THE GENETICS AND PHENOTYPING

OF

SEED MORPHOLOGY IN HEXAPLOID WHEAT

by
Keith Richard Williams
May 2013

A Dissertation
Presented to the Faculty of the Graduate School
of Cornell University
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy
in the Field of Plant Breeding

The Genetics and Phenotyping of Seed Morphology in Hexaploid Wheat

Keith Richard Williams, Ph.D.

Cornell University 2013

The size and shape of wheat seeds are important characteristics that are monitored in breeding programs due to the impact they may have on the two largest economic factors determining the success of a cultivar, quality and yield. Despite this importance, there are mixed reports on the ability of plant breeders to accurately phenotype as unique a characteristic as 'shape' or to effectively implement selection procedures based on the morphology of wheat seeds. Furthermore, the complex physiological and genetic relationships between individual components of seed morphology complicate the matter of effectively improving yield or quality via seed shape. Due to recent developments in both phenotyping procedures and methods of genetic analysis, plant breeders can now understand seed shape at the genotypic level and translate that knowledge into meaningful improvements to wheat breeding programs. In this dissertation, a thorough review of the literature surrounding seed shape in hexaploid wheat is provided, as well as results of experiments comparing different phenotyping methods, clarification of QTL underlying seed shape in both adapted and 'exotic' mapping populations, as well as preliminary work in phenotypic characterization of unique genetic resources for future mapping work using association mapping. This work adds to the current body of scientific knowledge by improving upon seed phenotyping methods, suggesting differential modes of action of QTL on specific dimensions of seeds, validating previously reported seed shape QTL, and showing the impact of divergent selection for seed fill period on kernel morphology. It is hoped this work will facilitate the further improvement of not only wheat, but all plant organs of economic importance.

BIOGRAPHICAL SKETCH

The author was born in Southeast Michigan, into the sprawl surrounding the remnants of Detroit. He was fortunate to be raised by free spirits, marry the love of his life, and discover a creative passion for plants.

He has benefited from knowing wonderful, eccentric people and looks forward to meeting many more in the future.

In his free time he can be found working in his garden or foraging for wild mushrooms.

I found a dimpled spider, fat and white,
On a white heal-all, holding up a moth
Like a white piece of rigid satin cloth --
Assorted characters of death and blight
Mixed ready to begin the morning right,
Like the ingredients of a witches' broth --
A snow-drop spider, a flower like a froth,
And dead wings carried like a paper kite.

What had that flower to do with being white,
The wayside blue and innocent heal-all?
What brought the kindred spider to that height,
Then steered the white moth thither in the night?
What but design of darkness to appall?—
If design govern in a thing so small.

-Frost

ACKNOWLEDGMENTS

Words can only partially express my deep gratitude to the following people who have contributed to this dissertation:

The Cornell Department of Plant Breeding,
for providing funding and a creative, dynamic environment to study.

Mark Sorrells, for patient guidance in teaching me the science of plant breeding.

Gary Bergstrom, for his cheerful discussions and whole-hearted inclusion of me in his lab group.

David BenDaniel, for helpful advice and encouragement of my entrepreneurial spirit.

Walter DeJong, for listening and support.

My wife, Amanda, who carried me when I felt broken.

My parents, who continually encourage me to pursue what I love.

My close friends in Ithaca, who made the weekends fun.

Donna Papatheodoropolous, Bruce Parfitt, Tracy Wacker, Ben Jones, and Mike Kearsey; all were amazing mentors that helped me along the way to Ithaca.

David Benscher, James Tanaka, Gretchen Salm, John Schiffer, Roxanne VanWormer, Jesse Munkvold, Mahmmoud Zeid, as well as many other students and visiting scientists in the Cornell Small Grains Breeding Project – for their thoughtful comments on my work, stimulating conversations, and occasional off-color humor.

TABLE OF CONTENTS

CHAPTER ONE

The genetics of kernel morphology in hexaploid wheat: changing technologies and their role in understanding wheat seed size and shape

I. Introduction

The wheat kernel has traveled with humankind over 10,000 years, giving rise to civilizations and serving as the foundation for numerous foodstuffs. Common wheat (*Triticum aestivum* L.) has been central to agriculture and the subject of much scientific study. While significant gains have been made in understanding the physiology and genetics of the wheat plant, there are still aspects of its biology that remain unclear. Among these unknowns are the genetic factors that regulate grain formation and kernel morphology. This lack of knowledge has hindered effective manipulation of seed morphology for improvement of wheat cultivars.

Seed biology has been investigated in numerous plant species including wild, domesticated, and model plants. Genes regulating key processes in seed formation such as size and dispersal mechanisms have received attention in crops such as rice (*Oryza sativa*) (Li *et al.* 2004; Fan *et al.* 2006; Guo *et al.* 2009; Takano-Kai *et al.* 2009; Yan *et al.* 2009), maize (*Zea mays)* (Li *et al.* 2010), and barley (*Hordeum vulgare*) (Mather *et al.* 1997; Igartua *et al.* 2000; Ayoub *et al.* 2002). The genetics of seed size have also been explored in model species such as *Arabidopsis thaliana* (Alonso-Blanco *et al.* 1999; Luo *et al.* 2005; Berger *et al.* 2007). In crops such as rice, knowledge of domestication-related phenotypes has been extended to studies of gene function (Lu *et al.* 1996; Doganlar *et al.* 2000; Peng *et al.* 2003; Willcox 2004; Li *et al.* 2004; Li *et al.* 2006). In wheat however, no studies have confirmed the structure or function of genes which specifically determine seed morphology.

Despite the current lack of knowledge about the genes affecting seed formation in wheat, directed manipulation of seed size and shape is a goal of wheat breeding programs due to the impact it

may have on other traits. Two primary considerations are given to wheat cultivars under consideration for commercial release: yield and quality. Size and shape of wheat kernels are potentially able to drive future gains in yield as well as quality. Therefore manipulation of seed dimensions could be used to identify wheat germplasm with greater potential. However, complex physiological relationships between seed morphology and other characteristics have hindered effective selection for yield or quality based on phenotypic selection of seed dimensions. Due to recent advances in genetic mapping methods, phenotyping technology, as well as greater knowledge about the genomes of related grass species, researchers are now poised to extend more thorough genetic studies of seed characteristics to hexaploid wheat.

II. Physiological Aspects of Kernel Morphology

The dimensions of a seed are closely related to other physiological processes occurring within the plant. Flexible variation in seed size, along with variation for germination and other seedling traits, helps plants compete in natural ecosystems. In wild species such variation is important for survival. However, the physiological interaction of seed characteristics can be disadvantageous to domesticated wheat cultivars. The challenge of plant breeding is modifying organisms that evolved in a competitive natural context to maximize their productivity in a very different, largely artificial system. The interaction of multiple seed characteristics in domesticated wheat, such as compensation between number of seeds and size of seeds, can be viewed as a remnant from survival mechanisms in wild ancestors. It is difficult to manipulate such interacting characteristics using phenotypic selection alone.

The physiological interaction between various reproductive characteristics, including seed morphology traits, is a survival mechanism meant to cope with source limitations in a natural environment. By examining a wide range of plant species, evolutionary models have been proposed to explain the

antagonistic relationship between seed size and seed number in plants. These models attempt to determine an ideal balance between the size and number of seeds for survival of the organism in a constrained environment. Multiple factors are considered in such studies, including size effects on dispersal, perennial lifestyle, competitive ability, and protection against sporadic environmental fluctuations. Genetic variability of seed size is described as a mechanism to cope with these events in wild species, but may be seen as an undesirable remnant of a crop's wild progenitors. This genetic variability is the result of alleles that are disadvantageous in modern agriculture.

The negative effect of alleles favoring flexible variation in seed characteristics is of concern in an annual crop that has the benefit of highly developed agricultural systems. Agriculture is different from nature; by definition, source limitations are not as prominent in agricultural systems. As a result of this, sink limitations may constrain production. Because modern agricultural practices remove natural resource limitations traits governed by sink capabilities, such as seed size, may be sub-optimized. In wild populations, Geritz *et al.* (1998) argues that competition among seedlings favors the evolution of variation in seed size within as well as between individual plants. In the context of a competitive natural environment, the relationship is necessary due to limited resources. Such an environment contrasts with agricultural systems. In agricultural systems, studies have found seed size in wheat to be sink limited (Brocklehurst 1977; Chojecki *et al.* 1983; Slafer and Savin, 1994). Since seed size is an important component of yield in wheat and is sink limited, understanding which regions of the genome regulate seed size is a goal of plant breeders. The molecular basis of phenotypic variation in wheat and how it can impact breeding programs is well documented (Gupta *et al.* 2010). Better characterization at the molecular level coupled with genotypic data could then be used to aid selection for genes that influence seed morphology and yield. Such information can contribute basic knowledge about what types of factors regulate resource partitioning in wheat and provide information about allelic diversity for this characteristic within breeding germplasm. More practically, mapping studies targeting seed

characteristics contribute useful genetic markers for selection of alleles that favorably impact sink-limited yield components such as seed size and morphology.

Seed size is determined by a number of physiological and genetic factors, including resource partitioning dynamics and the effects of major adaptation genes (e.g., vernalization, reduced height, photoperiod sensitivity) (Evers 1971; Briarty *et al.* 1979; Snape *et al.* 1985). Final grain size is also affected by heritable characteristics including seed vasculature, endosperm cell expansion, endosperm cell number, and the grain filling period (Cochrane 1983; Lingle and Chevalier 1985; Ugalde and Jenner 1990a; Ugalde and Jenner 1990b; Mou and Kronstad 1994a; Mou and Kronstad 1994b; Drea *et al.* 2005). These characteristics encompass the rate and duration of starch synthesis as well as the formation of starch/protein complexes in the developing grain (Chojecki *et al.* 1983). Additionally, characteristics such as glume architecture or light penetration into the floral cavity may play a significant role in regulating grain growth (Engledow 1920; Lamba 1949; Radley 1981; Millet and Pinthus 1984; Millet 1986; Raju and Srinivas 1991). Selection for larger seeds, or for increase in any specific dimension, is complicated by these interrelated pathways affecting seed development. Phenotypic plasticity, or the dynamic interaction of individual component traits in response to the environment, makes selection difficult for traits that physiologically interact. Recognizing the interrelated nature of such traits and seed shape is important when attempting to modify kernel morphology. Identification of the individual genes underlying seed characteristics per se does not solve the breeder's task of improving yield or quality via seed shape. To achieve meaningful gain from selection, a breeder must select for alleles that modify seed shape without negative pleiotropic effect (or perhaps with beneficial pleiotropic effect) on the interrelated downstream characteristics. Because of this, improved genetic information on seed characteristics must be used in conjunction with knowledge of the physiology of seed development to be useful to breeding programs.

Despite the relationships between many seed characteristics, not all regions of the genome controlling these traits are pleiotropic. Moreover, regions with very specific effects on individual seed characteristics can be identified using molecular markers. Some genes controlling the phenotypic interplay between seed size and number are pleiotropic due to their physiological interaction, but quantitative trait loci (QTL) studies have shown that closely related seed characteristics can be independently controlled at the genetic level. Among these independently inherited characteristics are the dimensions of seeds; studies of wheat kernel shape have shown length to be governed by distinct regions of the genome separate from those of width (Chojecki *et al.* 1983; Breseghello and Sorrells 2007; Sun *et al.* 2009). Seed morphology QTL have relationships with other reproductive traits such as grain number; while some of these interactions are antagonistic, alleles can be identified that affect size without changing kernel number (Chojecki *et al.* 1983). Such evidence suggests that failures to simultaneously select for seemingly antagonistic traits may be due to the non-specific selection among multiple QTL or genes when using phenotypic data alone. For example, if several QTL impact seed size, but only one is able to increase seed size and not have a negative effect on seed number, a researcher using only large seeds as a phenotypic selection criterion has no way of specifically selecting for specific alleles which increase size and do not reduce seed number. As a result of non-differential selection among several sites in the genome, selection is rendered less effective and the resultant phenotype appears to be canalized when in actuality it may not be. This problem is due to insufficient phenotyping (examining only a very limited number of seed traits) or lack of molecular marker information for use in conjunction with more traditional phenotyping approaches.

Genetic regions altering phenotype can be identified and clearly linked to particular grain shape components by measuring kernel shape in specific dimensions. The QTL underlying shape components can then be compared to other grain characteristics (thousand kernel weight, seed number, quality parameters) to clarify specific allelic relationships. Instead of reporting only on a single phenotype such

as thousand kernel weight (TKW), research groups are now phenotyping multiple seed development characteristics despite the highly intensive work required (Zhang *et al.* 2010). This approach identifies regions of interest within the genome and also clarifies pleiotropic or unique effects of these regions based on comparison. Identifying sites in the genome which independently affect seed dimensions allows use of molecular markers to build refined haplotypes that would be difficult to obtain through phenotypic selection measures alone. In this manner, the use of genotypic information would circumvent the difficulties of phenotypic selection encountered due to phenotypic plasticity.

III. Economic Aspects of Kernel Morphology: Yield

Yield of wheat is conditioned by the size and shape of individual seeds but these individual characteristics have not been used to drive increases in wheat yields. Yield is the product of a number of physiological processes occurring in the life cycle of the plant, influenced by their interactions with the environment. Plant breeders have succeeded in increasing yields in wheat, but it has not been the result of targeting seed size or shape. In wheat, the failure of improving yield through directed manipulation of kernel morphology can be clearly defined.

Yield of wheat is a product of the number of inflorescences per plant at a specific plant density, the number of kernels per inflorescence, and the size of the individual kernels. All of these factors are interrelated physiologically and respond to environmental cues (Del Moral *et al.* 2003). Of the individual yield components, seed size is the more stable than seed number across environments (Petrovic and Worland 1988; Giura and Saulescu 1996; Peltonen-Sainio *et al.* 2007). Because seed size is more stable, it makes an easier target for manipulation in order to increase yield.

While researchers have devoted much time to describing and examining the different physiological components that contribute to yield, plant breeders typically use direct yield data rather than yield component values when it comes to making selections. Using phenotypic selection of yield component traits to effectively improve yield remains poorly understood and debatable to wheat breeders (Yamazaki and Briggle 1969; Hook 1984; Parker *et al.* 1999). Empirically, Wiersma *et al.* (2001) focused on selecting for larger (i.e. greater length & width) seeds and while they were able to improve quality characteristics, yield was unaffected (Wiersma *et al.* 2001). Concurrently, number of kernels per spike and number of tillers decreased. Thus, phenotypic selection based on increasing seed size alone did not improve yield due to compensation by other yield components. This illustrates the complex physiological relationships that make it difficult to manipulate yield using component traits. Sayre *et al.* (1997) found little evidence to support a direct relationship between yield increases and changes to the number or size of seeds on an individual plant basis. In their retrospective study of phenotypic changes to wheat using 30 years of breeding data from the International Maize and Wheat Improvement Center (CIMMYT), there was no correlation observed between yield increase and changes in seed size. Lack of correlation between yield increases and seed size suggests breeding efforts have not been utilizing seed morphology to drive yield increases, but rather that researchers have been manipulating other factors contributing to yield gains. If the antagonistic physiological relationships among individual seed characteristics cannot be resolved using current phenotypic methods, selection of large seeds with the goal of improving yield will predictably fail.

That breeders have been ineffective in manipulating seed morphology to help increase yield in wheat can be seen as the result of two major factors: 1) Phenotyping yield components is time-consuming and expensive. Unless there can be justification for collecting data on yield components that extends to characteristics other than yield alone, it is more cost effective to use harvest data. 2) Even if rapid phenotyping can be adopted for yield components using widely available tools such as sizing screens or

seed counters, there is often a failure of phenotypic selection based on these procedures due to the physiological trade-off between yield components. The implication of these two observations is that breeding programs must actively pursue and evaluate more cost-effective, precise phenotyping methods in conjunction with the use of genotypic data in order to effectively increase yield by manipulating seed size or shape. This becomes increasingly important as breeding programs accelerate the pace of breeding through reductions in cycle time using off-season nurseries, doubled-haploids, and phenotypic prediction methodologies (Peleman and van der Voort 2003; Eathington *et al.* 2007; Heffner *et al.* 2010).

IV. Economic Aspects of Kernel Morphology: Quality

Milling quality is important to determine the value of a wheat cultivar. This characteristic is tested using small scale milling evaluations. These evaluations are expensive and time consuming, limiting their use. As a result, wheat breeders have sought to relate milling quality to phenotypes that are cheaper or easier to evaluate. The relationship between milling yield and the individual dimensions of kernels has been proposed as a way to achieve this goal. By reviewing relevant background literature Marshall *et al.* (1984) attempted to find a relationship between kernel dimensions and milling yields. The work from that study and provided a thorough review of both theoretical aspects and empirical data. They proposed that better milling yields could be achieved by actively selecting for a larger, more spherical wheat kernel. Seeds with spherical shape of larger volume should have increased endosperm content and reduced surface area (bran). This increased endosperm content would presumably increase the flour yield. However, the empirical results from Marhsall *et al.* (1986) did not support theoretical predictions. The discrepancy between the empirical and theoretical studies was postulated to be due to a number of other factors that must affect milling quality, such as linkage drag (Marshall *et al.* 1986). A

lack of relationship between kernel morphology and milling qualities has also been reported by other research groups (Bergman *et al.* 2000; Schuler *et al.* 1995). These studies all examined the subject using phenotypic data alone.

Despite these discouraging results, positive relationships between grain size (typically measured as TKW) and flour yield have been reported as well. These relationships have extended across multiple populations using phenotypic surveys, controlled-cross QTL mapping studies, and more recent association analysis techniques (Berman *et al.* 1996; Baker *et al.* 1999; Campbell *et al.* 1999; Novaro *et al.* 2001; Wiersma *et al.* 2001; Breseghello and Sorrells 2006). Notably, the SSR marker *Xbarc232-5b* was demonstrated to have significant association with milling traits and seed size among a panel constructed to represent the diversity in quality of northeastern U.S. soft winter wheat cultivars (Breseghello and Sorrells 2006). Such reports suggest that the ability of molecular marker approaches targeting seed characteristics may aid in improving wheat cultivar quality. It is expected that such results could be supported by other studies specific to breeding materials of interest. These could then form the basis for improving milling quality in target germplasm via selection based on kernel shape in conjunction with genotypic data.

V. Genetic Aspects of Kernel Morphology

Genetic architecture of seed size and shape in wheat

Numerous studies have tried to understand the genetic determinants of seed characteristics in wheat. Early studies were conducted using cytological mutants to identify which chromosomes, or segments of chromosomes, had large effects on seed size (Law 1967; Halloran 1976; Snape *et al.* 1985; Backes *et al.* 1995). Even with the hindrance of only being able to confirm the role of large portions of chromosomes, studies using chromosomal deletion mutants were successful in determining the quantitative nature of seed size and the wide dispersal of relevant genes across the genome.

Genetic studies using controlled-cross QTL mapping methods supported chromosome deletion experiments, and found that multiple regions of the genome harbor QTL affecting grain weight (Giura and Saulescu 1996; Varshney *et al.* 2000; Peng *et al.* 2003; Quarrie *et al.* 2005). In these studies, measures of grain weight such as TKW were used to describe seed size. QTL affecting seed size can be found across all chromosomes of wheat, with varying degrees of effect seen for individual QTL (Campbell *et al.* 1999; Dholakia *et al.* 2003; Breseghello *et al.* 2005; Quarrie *et al.* 2005; Huang *et al.* 2006; Sun *et al.* 2009; Gegas *et al.* 2010; Tsilo *et al.* 2010). A recent meta-QTL study has compiled the results of many of these and identified regions on 1A, 1B, 2A, 2D, 3B, 4A, 4B, 4D and 5A that are frequently cited as influencing seed morphology (Zhang *et al.* 2010).

QTL studies can locate the general regions of genes of interest but do not have a high enough resolution in wheat to allow mapping via a candidate gene approach. Confidence intervals for seed size QTL can range from 1cM to more than 30cM, with many QTL studies returning confidence intervals around 10cM. Given a 16,000 Mb genome (Aragumuganathan and Earle 1991) and a total genetic map length of 2,569 cM (Somers *et al.* 2004), 10cM genetic distance would on average represent a physical distance greater than 6,000kb. It is known that genes are not evenly distributed across the wheat genome (Sandhu and Gill 2002). Even under a simplified (and erroneous) assumption of even gene distribution, candidate gene approaches to cloning are difficult when the average indications using biparental mapping populations cannot narrow the focus to physical regions of less than 6,000kb. Coupled with a genome that is low in polymorphism and far from being sequenced, accurately locating or characterizing genes in wheat is a challenging prospect.

Heritability and environmental influence on seed size and shape

Seed size and shape have moderate heritability and are less environmentally influenced than other yield component traits. TKW is a normally distributed trait. The length, width, and thickness dimensions of seeds describing their shape also show a normal distribution (Breseghello and Sorrells 2007). Grain size is less affected by environmental effects than grain number per ear; grain number per ear is described as grain set and is determined prior to and during anthesis (Petrovic and Worland 1988). Because the success of grain set is determined over a small time frame during the season, the effect of sporadic environmental stresses such as drought are likely to have a greater effect on the number of kernels per ear than on the size of those kernels (Del Moral *et al.* 2003). Kernel characteristics have a moderate to high heritability, with size (as TKW) ranging from broad-sense heritability of 0.58 to 0.90 and shape parameters (length, width) ranging from 0.55 to 0.95 (Barnard *et al.* 2002; Sun *et al.* 2009; Wang *et al.* 2009; Gegas *et al.* 2010; Tsilo *et al.* 2010). In general, the trend seems to be that the TKW of cultivars is more heritable than shape parameters and length of kernels is more highly heritable than width (Sun *et al.* 2009). From a physiological standpoint, the observation of differential heritability fits evidence supporting sequential development of yield components (Kozak and Madry 2006). Length of a seed is set earlier in the developmental process whereas the width of a seed has more time to be influenced by environmental conditions during the seed filling period (Sadras and Egli 2008). Seed morphology makes a more logical target for breeding programs than other yield components such as seed number per inflorescence because the lesser influence of the environment and moderate (0.55 – 0.90) heritability of seed dimensions.

Known genes with qualitative effects on seed morphology

Major genes influencing adaptation of wheat cultivars are pleiotropic and also affect seed characteristics. Most prominent among these are the *rht* dwarfing genes, *ppd* genes influencing

photoperiod sensitivity (Scarth 1985), *vrn* genes for vernalization requirement, and genes that strongly influence multiple adaptive characteristics, such as the *S1/S2/S3* genes from wheat relative *Triticum sphaerococcum* and the *Q* gene (Snape *et al.* 1985; Salina *et al.* 2000; Snape *et al.* 2007).

Each of these major genes affect seed characteristics through varied physiological interactions. Genes controlling protein content of grain and specific enzymes involved in seed formation, such as Serpins, have been implicated in seed phenotypes (Rosenkrands *et al.* 1994; Rasmussen *et al.* 1996; Roberts *et al.* 2003; Cane *et al.* 2008). The *Q* gene on chromosome 5A determines the 'speltoid' phenotype of hexaploid wheat and influences seed size (Simons *et al.* 2006). Cultivated wheat has been selected for specific *Q* alleles along with several other genes conferring good agronomic type (Sourdille *et al.* 2000). Gene-specific markers have been developed for Q (Kato *et al.* 1999; Kato *et al.* 2003; Simons *et al.* 2006; Asakura *et al.* 2009; Takano-Kai *et al.* 2009). The dwarfing genes *rht* have been studied thoroughly for their pleiotropic effect on a number of other traits in wheat, including yield and yield components (Fischer and Stockman 1986; Keyes 1989; Fischer and Quail 1990; Ellis *et al.* 2004; Rebetzke *et al.* 2000; Rebetzke *et al.* 2012). Flintham *et al.* (1997) found that, depending on the background in which the various *rht* alleles were deployed, there were varying effects on the relationship between the number and size of wheat kernels. Almost universally, introgression lines carrying *rht* alleles produced more kernels. However, some lines compensated by producing smaller individual kernels whereas others produced more kernels without a reduction in individual kernel size, thereby increasing yield (Flintham *et al.* 1997). This observation supports the genetic evidence that interacting yield components can be regulated independently. This also suggests that greater knowledge of allelic diversity coupled with molecular markers would allow concurrent selection for both larger and more numerous kernels.

The presence of major genes complicates the identification of other unique genes controlling seed characteristics due to the pleiotropic effects that major genes may have on seed characteristics. Many

chromosomes in wheat harbor a number of major genes influencing phenotype, including awnedness-inhibitor *B1*, *rht8*, *rht12*, *vrn1*, and *Q* (Korzun *et al.* 1997; Korzun *et al.* 1998; Kato *et al.* 1998; Kato *et al.* 2000; Kato *et al.* 2003). Both dwarfing genes *rht8* on 2D and *rht12* on 5A are in regions identified as influencing kernel morphology in numerous QTL and association mapping studies. In addition to *rht8*, *ppd1*, a major gene influencing photoperiod sensitivity, is located close to regions influencing seed morphology on homelogous group 2 (Mohler *et al.* 2004; Goncharov and Watanabe 2005). The serpin gene (*Ser5B*), grain-protein content (*Gpc-B1*) gene, and the *S* genes may also influence seed characteristics of wheat. *Ser5B* encodes a serine proteinase inhibitor on chromosome 5B that has been suggested to have a defensive role in preventing insect damage or oxidative stress, as well as perhaps influencing grain quality traits due to similarities between serpin protein motifs and storage proteins in the endosperm (Cane *et al.* 2008). Of 150 proteins surveyed during grain development, serpin was one of only 17 found to be increasingly expressed throughout the grain development period (Nadaud *et al.* 2010). The *Gpc-B1* gene on chromosome 6B encodes a transcription factor which accelerates senescence and affects a number of other traits, including grain size (Uauy *et al.* 2006; Distelfield *et al.* 2006; Waters *et al.* 2009). Yet another collection of major genes with pleiotropic effects are the *S* genes located at three sites dispersed among the three homeologous genomes (3A, 3B, 3D) within wheat (Maystrenko *et al.* 1998; Salina *et al.* 2000). The *S* genes are mutations found in 'shot wheat' (*Triticum sphaerococcum*) which induce round seed shape and affect other adaptive traits (Salina *et al.* 2000). Whether the QTL detected for seed morphology near these chromosome regions containing major genes are due to their effects or unique alleles is unknown.

Quantitative traits having physiological relationships with seed morphology

In addition to the effects of major qualitative genes, a number of quantitative traits affect the final size of wheat seeds. The major dimensions of the seed, its symmetry as well as surface texture, and the grain filling dynamics of a cultivar may influence seed morphology. A wheat kernel can be roughly described as an ellipsoid whose volume is defined by the axes corresponding to length, width, and thickness. Based on a strictly geometric definition, a wheat kernel would be expected to have a greater overall yield if seeds were larger and seed set remained constant. Greater flour yield would be expected as seeds increased in size and became more spherical. A more spherical seed would maximize flour yield by having the greatest ratio of endosperm (flour) to seed coat (bran) (Marshall *et al.* 1984). In the same way, the more subtle aspects of shape such as asymmetry or perimeter roughness contribute to how spherical a kernel is and its potential flour yield. Relating seed shape to flour yield would be valuable for breeding programs, as current small-scale flour yield assessment procedures are costly and time-consuming (Finney and Andrews 1986; Andrews 2002).

QTL studies have identified regions of the genome associated with individual kernel characteristics. Results from controlled-cross QTL studies have found suspected pleiotropic loci that affect multiple kernel traits; markers associated with these include *Xgwm261* (2D) influencing length, width, and weight, and *Xgwm515* (2D), which influenced length, factor form density, and weight (Dholakia *et al.* 2003). Some alleles are identifiable only in particular locations while others are stably detected across environments (Breseghello *et al.* 2005). The negative effects of alleles on kernel size suggest that even large-grained cultivars may have room for improvement through elimination of alleles negatively affecting kernel size (Giura and Saulescu 1996).

Furthermore, variation for these genes could be mined from wild relatives through the use of molecular markers. A crop such as wheat that has undergone two inter-taxa hybridization events would have less

chance to carry over diverse alleles than a diploid crop that humans could repeatedly integrate into their fields during the domestication process. Presumably there would be diverse alleles for quantitative traits that would have been left behind at these population bottlenecks. Recent studies provide evidence supporting the existence of alleles in non-adapted germplasm with positive effects on agronomic characteristics (Gegas *et al.* 2010). That seed morphology in domesticated wheat could contribute to improved cultivars and that variation still exists in wild relatives indicates that future yield increases could be driven by manipulating seed characteristics using both cultivated and wild materials.

Other quantitative traits determining adaptation of a cultivar are related to seed characteristics. The grain-filling period of a cultivar can affect seed size and shape. Well-adapted wheat varieties can take advantage of the full growing season and develop during the entire period of available favorable growing days. In wheat, a longer filling period generally produces larger grain but runs the risk of being affected by biotic or environmental stress (Motzo *et al.* 2010). Grain filling is also affected by floret position in the wheat spike, with the centermost florets filling first and typically producing the largest grains (Keefe 1990). Recent QTL studies have examined the rate and length of fill period and found that there is co-localization of some of these regions with those conferring grain sizes (Wang *et al.* 2009). However, there remain independent sites for determination of filling period characteristics and kernel morphology (Wang *et al.* 2009). Some QTL affecting filling characteristics are independent of those affecting seed size. This is important, as breeders will ideally want to identify genetic targets for manipulation of seed size that will not disrupt a cultivar's adaptation to particular growing regions.

VI. The Role of Changing Technology

New Genetic Mapping Approaches – Association Mapping

Association mapping (AM) was developed in animal breeding and human genetics studies, where large biparental mapping populations are not feasible. Biological constraints have required geneticists from these fields to develop mapping approaches constructed under the assumption of large numbers of individuals in random mating populations. Similar to traditional linkage analysis, the theory behind AM is that any two sites on a particular linkage group will travel with each other over generations until these two sites are separated by a recombination event during meiosis. As a result, more distant sites will have a greater probability of a recombination occurring between them. Therefore, over multiple generations, closely linked loci will tend to be observed together. Simultaneous observation of markers with particular phenotypes then allows identification of markers that are in tight physical linkage to genes of interest via correlation. Relying on past meiotic events rather than newly generated recombination (as in controlled-cross mapping) is what largely differentiates AM from traditional QTL mapping approaches.

Fine mapping of genes using AM in plant populations has great potential where traditional QTL approaches have failed. While the specific mapping resolution of any biparental population is unique to population size and varies across the genome, controlled-cross QTL mapping studies in wheat often detect QTL with large confidence intervals. The ability to improve mapping resolution is particularly relevant in wheat, where these large confidence intervals can encompass a huge amount of sequence content. Because AM is based on the statistical association of loci using numerous historical recombination events, it allows mapping of genes at a higher resolution than in controlled-cross QTL studies where there are limitations to the number of informative recombination events. Association mapping offers other advantages for genetic studies such as the ability to map using panels of distantly

related lines. Because of this, AM offers a way to map genes of interest in collections of breeding materials with diverse phenotypes thus circumventing the need to develop multiple new mapping populations for genetic studies and divert resources from breeding activities. Since AM panels are composed of loosely related individuals, they have another advantage over traditional mapping populations in that AM panels sample a wide range of alleles (Flint-Garcia *et al.* 2003; Breseghello and Sorrells 2006; Rafalski 2010). Thoughtful construction of such panels allows for the development of a resource that may be shared among multiple research groups. An excellent example of the benefits of such an approach in plant breeding can be seen through the work being conducted using the maize nested association mapping panel (NAM) (Yu *et al.* 2008).

Association Mapping - challenges of population structure in plants

Until the last decade, the presence of population structure created through artificial selection and inbreeding has been a major barrier to the widespread use of AM in plants. The theory of AM is constructed under the assumption of large numbers of individuals in random mating populations. This assumption is incorrect in many crop species, particularly if the species are self-pollinated. Frequently a limited number of founder parents are used to generate many progeny lines that can be carried over multiple generations and are recycled back into crosses in a breeding program. Effectively, the small number of successful genotypes that give rise to large plant breeding populations creates high levels of kinship, or relatedness, between lines derived from particular families. The structure within families, denoted as kinship and treated statistically as 'K', previously confounded the use of AM in plants. Early attempts to resolve kinship structure present in plant populations had been based on pedigree records using coefficient of coancestry (Falconer and Mackay 1996) – this was problematic because often pedigree records were inaccurate or missing.

However, with the widespread adoption of molecular markers to characterize plant genotypes, geneticists began to use marker information rather than pedigree records to calculate kinship between lines with better accuracy. Accurate calculations of kinship were achieved using mixed-model approaches, initially in maize (Yu *et al.* 2006). More recently such studies have been extended to compare other approaches such as restricted maximum-likelihood (REML) to account for kinship, as well as explore the effects of these estimations in crops having different reproductive strategies and resultant population structure (Rafalski 2010). Each of these methods varies slightly, but relies on the use of molecular markers to clarify the genetic relationship between individuals and remove error that those relationships may add to genetic analyses.

Computational advances have also provided accessible tools to account for population structure. As a result, AM studies using plant breeding germplasm have now become practical (Pritchard *et al.* 2000). Significant to the success of AM in plants has been the widespread availability of software tools such as TASSEL (http://www.maizegenetics.net/) that integrate improved algorithms for determining population structure using molecular marker data with a user-friendly graphical interface. TASSEL (along with other similar programs) has stimulated a number of studies elaborating the use of AM in crop plants including maize, barley, and wheat, as well as in *Arabidopsis* (Andersen *et al.* 2005; Breseghello and Sorrells 2006; Crossa *et al.* 2007; Yu *et al.* 2008; Brachi *et al.* 2010; Chan *et al.* 2010).

Association Mapping - patterns of linkage disequilibrium and marker density

The ability to determine population structure does not remove the other main challenge of implementing AM in crop species: understanding the pattern of linkage disequilibrium in their genomes. This pattern will determine the number of markers required to make valid inferences. Linkage disequilibrium (LD) is the phenomenon of sites being inherited together over time; two sites are not

observed in the state of random equilibrium that would be expected if they were completely unlinked. For outcrossing species with frequent opportunities to undergo recombination, patterns of physically linked sites are quickly disrupted through frequent meiotic events. These populations will require high marker densities but will presumably have markers that are identified as significant only when they are in close proximity to target genes. However, many crops are able to self-pollinate. Such inbreeding can create large spans of DNA which are fixed early on and have no opportunity to be 'reshuffled' via outcrossing. These breeding lines represent highly homozygous individuals that are closely related, and therefore have large linkage blocks. Such large, monomorphic linkage blocks complicate the use of associations between two sites along a chromosome to interpret distance. These populations will require fewer markers, but significant markers may not actually be close to target genes (Breseghello and Sorrells 2006).

Challenges of AM posed by inbreeding populations with large linkage blocks can be resolved using special populations having reduced linkage block size. In wheat, manually crossing individuals in a population to increase recombination rates and break apart large linkage blocks would be labor intensive. However, a number of sterility systems exist that can aid in forcing outcrossing. Notable among these are dominant male-sterile (DMS) populations, which can easily be forced to outcross by selectively collecting seed from sterile plants at every generation (Sorrells and Fritz 1982). These can be used in a number of breeding schemes, but are particularly well suited to reducing linkage block size in the construction of panels for AM. Comparative studies of LD in 5A chromosomal regions of DMS materials against those in traditional wheat cultivars have shown an increased rate of LD decay (<1cM) to support this (Breseghello and Sorrells 2006; Heffner *et al.* 2008). Such a high rate of LD decay indicates that more than one marker per cM would be needed to achieve a reasonable power of detection, but is favorable for mapping in confidence intervals defined by controlled-cross mapping studies (Thornsberry *et al.* 2001). Combined with continually improving approaches for defining

significant associations in structured plant populations (mixed-model approaches) and increasingly thorough genotypic data, AM may be a way to close in on areas of interest in the wheat genome.

New Phenotyping Methods – quantitative measurement using digital image analysis

Precision of QTL analyses is also affected by variation in phenotyping methodology among studies. For describing seed size, TKW is commonly used because of the ease of measurement. While indirect measures of the shape or volume of seeds such as TKW are convenient, TKW encompasses a number of variables that are not fully defined and adds confusion when comparing results among experiments. For example, TKW could be a poor estimator for seed size when the actual volume of kernels is not considered. This is because it does not account the factor of density explicitly – two cultivars could have a similar TKW value but different volume/density relationships. To complicate the matter, there is a range of sampling methods used to calculate TKW. Some research groups count exactly 1,000 kernels (D. Benscher, Cornell Small Grains Breeding Program, personal communication); others have used multiple samples of smaller numbers and then converted to TKW (Bergman *et al.* 2000; Groos *et al.* 2003; Li *et al.* 2007; Sun *et al.* 2009). Rather than reporting these phenotypes as an estimate of weight of an individual seed in the context of their sampling methodology, researchers continue to convert to TKW. It has been pointed out that QTL studies will be more useful if phenotypes are broken down into component parts as far as possible and phenotyping can be made more precise (Paterson *et al.* 1991). Currently, technology is available that allows studies to move to more precise measurements of characteristics such as seed dimensions. Such technology that allows more precise and accurate phenotyping is being aggressively pursued by plant breeders as a way to take better advantage of low-cost genotyping resources that are increasingly available (Montes *et al.* 2007; DeSouza 2010; Houle *et al.* 2010; Poland *et al.* 2012).

Methods of measuring kernel shape have varied over the years, with early studies utilizing labor-intensive direct measurements of grain samples with calipers or rulers (Giura and Saulescu 1996; Dholakia *et al.* 2003). Increasingly, digital image analysis (DIA) is being used due to its speed and accessibility (Cober *et al.* 1997; Campbell *et al.* 1999; Diao *et al.* 1999; Horgan 2001; Ibaraki and Kenji 2001; Shouche *et al.* 2001; Doehlert *et a.l* 2004; Kwack *et al.* 2005; Breseghello and Sorrells 2006; Doehlert *et al.* 2006; Shimoji *et al.* 2006; Breseghello and Sorrells 2007; Dana and Ivo 2008; Himstedt *et al.* 2009). Also known as photometrics, the ability to convert images of plants to large data sets describing their dimensions is highly desirable for high-throughput phenotyping applications. When measuring kernel morphology such approaches provide a more direct approximation of kernel volume than TKW and allow examination of aspects of shape beyond gross individual seed dimensions. Some limitations of DIA have been identified previously (Tappan *et al.* 1987).

While length, width, and thickness remain important descriptors of plant organs, such as seeds, more sophisticated approaches to describing the highly variable shape of objects have been developed (White *et al.* 1988; Iwata *et al.* 1998). Although these methods vary, an increasingly popular approach is to use elliptical Fourier descriptors (EFDs) of an object's shape, followed by principle component analysis to reduce the highly complex EFD description to principle component values suitable for use in genetic analysis (Iwata *et al.* 1998; Ohsawa *et al.* 1998; Goto *et al.* 2005; Iwata *et al.* 2010). The use of EFDs allows for a quantitative measure of an object's shape beyond major geometric axes, area, and perimeter measures. Because of this, EFDs may provide useful phenotypic descriptions of surface texture, asymmetric size distribution, or other cryptic aspects of a seed's shape that cannot be described using the assumption of a general 'ellipsoid' shape. A freely available and easy-to-use program called SHAPE has been developed to facilitate generating such descriptions of plant organs (Iwata and Ukai 2002).

Another advantage of using DIA to phenotype kernel morphology is that once seed samples have been imaged it is easy to save image files for future analyses. The ability to reduce seed inventories and shift the workload of phenotyping from an immediate task to one that is more portable physically and temporally is very valuable.

New Information – comparative genomics of seed morphology in the grasses

In addition to technological advances in phenotyping, understanding of the genetics of other grasses can help identify genes influencing kernel morphology in wheat. Studies in rice have examined genes related to seed size and shattering as significant factors in domestication (Konishi *et al.* 2006; Li *et al.* 2006). The rice genome sequence and extensive marker resources have facilitated discovery and characterization of major genes affecting grain size. A grain-weight QTL has been localized to the pericentromeric region of rice chromosome 3, and mapped to a 93.8-kb region of the genome (Li *et al.* 2004), and grain-weight genes have been cloned and characterized (Fan *et al.* 2006; Agrama *et al.* 2007; Cho *et al.* 2007; Matsuoka and Ashikari 2007; Song *et al.* 2007; Xie *et al.* 2008; Guo *et al.* 2009; Takano-Kai *et al.* 2009; Yan *et al.* 2009). Knowledge of genes affecting seed characteristics in rice can provide preliminary genetic information for other grass species. Such comparative genomics approaches have been used to locate orthologs of genes affecting seed size of rice in maize (Li *et al.* 2010) and sequence similarity as well as chromosomal segment colinearity exists between rice and wheat (Wilson *et al.* 1999; Sorrells *et al.* 2003). Presumably information from rice seed size genes can be used to locate orthologs in wheat also.

Similarly, barley has been promoted as a model for grains due to its smaller diploid genome (Snape *et al.* 1996). Comparison of genomic regions between barley and wheat could help select important regions of the wheat genome for investigation using association genetics approaches. Studies of barley have found

QTL associated with kernel size and uniformity (Mather *et al.* 1997; Ayoub *et al.* 2002; Li *et al.* 2005). Barley studies have also located QTL that are associated with kernel plumpness and kernel weight (Mather *et al.* 1997). Some, but not all, of the QTL exhibited interactions with the environment, and these regions roughly corresponded to seed-related QTL identified in wheat mapping studies (Ayoub *et al.* 2002). Co-localization of regions influencing seed morphology in barley with QTL identified for seed morphology in wheat helps prioritize regions of interest for genetic studies.

VII. Improving Selection for Kernel Morphology

Quantitative Trait Ideotypes – haplotype construction using large-effect QTL

Marker-assisted selection (MAS) has found wide use in breeding for qualitative traits such as disease resistance (Gupta *et al.* 2010). While MAS may be unable to capture small-effect QTL, pyramiding QTL regions with large effect (or actual genes) on quantitative traits is a useful first step in manipulating these traits. Selection of QTL with large r^2 values can function as the first step in a two-tiered approach, where MAS pyramiding of large-effect QTL is followed by selection of small effect QTL through extensive replicated testing or genomic selection. Deeper understanding of which regions of the genome are responsible for governing the shapes of wheat kernels may be identified through AM studies and provide a defined target for breeders to implement such an approach. Following identification of markers in tight linkage with genomic regions of interest, breeders could work to produce cultivars with specific kernel dimensions, or optimized shape and size. While "optimized" is a tricky word in plant breeding, in this case it refers to a shape which can be subjected to selection and demonstrably contributes to greater yield and quality of the cultivar. It is important that alleles are identified in material that is as closely related to target germplasm as possible, because it has been shown that QTL

are often population specific and it may be difficult to effectively transfer to breeding programs from special mapping populations (Stuber *et al.* 1999).

The science of plant breeding has centered on the need for quantitative data and repeated evaluations of genotypes to support a breeder's choice in cultivar development. However, this is not to say that directed manipulation of plant form towards an ideal is an unwarranted goal in plant breeding programs. Indeed two of the major contributions to world agricultural productivity have been the development of dwarf wheat and rice varieties. These materials represent the successful application of an ideotype, or optimized plant form, based on sound reasoning that shorter stature plants would suffer less lodging and could be heavily fertilized. Development of such cultivars played a central role in the green revolution of the 1970s that is credited with boosting global yields (Hedden 2003; Pearce *et al.* 2011). Introduction of dwarf cultivars is a single example of the concept of ideotype, which has consistently reoccurred in plant breeding as geneticists try to engineer their concept of what constitutes a 'good' plant. Such directed effort to change plant morphology, architecture, or general form has typically been centered on genes with qualitative effects. Striving for particular grain shapes is no different in concept, but is more difficult to accomplish. Whereas development of dwarf wheat varieties was accomplished through use of dwarfing genes with qualitative effects on phenotype via backcrossing, kernel morphology is a quantitative trait that has lower heritability. Therefore, the challenge of building ideotypes around kernel shape in a breeding program will rely on improving marker robustness and implementation of selection methods utilizing well-defined phenotyping criteria.

Increasing Genetic Diversity – targeting seed morphology in pre-breeding germplasm

Marker assisted selection of seed characteristics could be used to increase allelic diversity in cultivated wheat germplasm. There is wide phenotypic diversity in seed morphology among wild relatives of

wheat (Gegas *et al.* 2010). Following wide crosses to such materials, pre-breeding work could be focused on selection for alleles conferring large seed size or favorable seed shape. Use of markers identified by AM could be used to increase diversity in adapted breeding pools via enrichment for rare alleles. A rare-allele enrichment approach could be especially useful with respect to seed morphology, as a number of newly created allohexaploid 'synthetic' wheats have been proposed as breeding materials due to their large seed size. By identifying alleles that positively affect seed shape without negative pleiotropic effects on adaptation or other characteristics, synthetic hexaploid wheat and wild relatives could be more effectively used to improve seed characteristics in cultivated wheat (Skovmand *et al.* 2001; Sorrells *et al.* 2011).

Over 30 years of genetic research have shown that multiple regions of the genome control the size and shape of wheat kernels but have made little practical impact in the way breeding programs are conducted. These studies vary in the populations they examine, the details of their phenotyping methods, and the amount of genotyping resources that were available at the time of their publication. Little of the information from these studies has been documented as having practical use in breeding programs, and gathering the information has required diversion of resources from more immediate breeding program activities. However, advances in phenotyping, genetic mapping, and knowledge of other grass species' genomes offer promising solutions to using individual kernel characteristics to improve the yield and quality of cultivars.

In summary, refined phenotyping assays for seed morphology may be used in conjunction with association mapping techniques to improve the efficiency of modern breeding programs. Improved genetic information about seed morphology could be used to identify valuable alleles for increasing seed size without negative pleiotropic effects on other seed characteristics. Use of DIA techniques for

exploring the genetics of organ shape in plants are amenable to modern breeding because DIA studies (1) are non-destructive phenotypic assays which can be used upstream in the breeding process to phenotype individual plants during early generations of selection or seed increase, (2) leverage advances in technology to assign quantitative values to traits that were formerly only able to be categorized, and (3) provide a phenotyping assay that is digitized, thus allowing distribution of the phenotyping process over time. Such characteristics are of prime importance as numbers of plants evaluated increase and available phenotyping time between reseeding decreases, and emphasize the need for more exploration of the use of these techniques in crops such as wheat.

REFERENCES

Agrama H, Eizenga G, Yan W (2007) Association mapping of yield and its components in rice cultivars. Mol. Breed. 19:341-356

Alonso-Blanco C, Blankestijn-de Vries H, Hanhart CJ,Koornneef M (1999) Natural allelic variation at seed size loci in relation to other life history traits of arabidopsis thaliana. Proc. Natl. Acad. Sci. U. S. A. 96:4710-4717

Andersen JR, Schrag T, Melchinger AE, Zein I, Lubberstedt T (2005) Validation of *Dwarf8* polymorphisms associated with flowering time in elite European inbred lines of maize (*Zea mays* L.). Theoretical and Applied Genetics 111:206-217

Andrews L (2002) Quality Characteristics of Soft Wheat Cultivars. USDA ARS Soft Wheat Quality Laboratory, Wooster, OH

Aragumuganathan K, Earle ED (1991) Nuclear DNA content of some important plant species. Plant Molecular Biology Reports 9:208-218.

Asakura N, Mori N, Nakamura C, Ohtsuka I (2009) Genotyping of the Q locus in wheat by a simple PCR-RFLP method. Genes Genet Syst 84:233-237

Ayoub M, Symons S, Edney M, Mather D (2002) QTLs affecting kernel size and shape in a two-rowed by six-rowed barley cross. Theoretical and Applied Genetics 105:237-247

Backes G, Graner A, Foroughi-Wehr B, Fischbeck G, Wenzel G, Jahoor A (1995) Localization of quantitative trait loci (QTL) for agronomic important characters by the use of a RFLP map in barley (*Hordeum vulgare* L.). Theoretical and Applied Genetics 90:294-302

Baker S, Herrman TJ, Loughin T (1999) Segregating hard red winter wheat into dough factor groups using single kernel measurements and whole grain protein analysis. Cereal Chemistry 76:884-889

Barnard AD, Labuschagne MT, van Niekerk HA (2002) Heritability estimates of bread wheat quality traits in the Western Cape province of South Africa. Euphytica 127:115-122

Berger F, Gerald JF, Ingouff M (2007) *Arabidopsis* as a model for understanding the basics of endosperm development. Endosperm 8:91-110

Bergman CJ, Gualberto DG, Campbell KG, Sorrells ME, Finney PL (2000) Kernel morphology variation in a population derived from a soft by hard wheat cross and associations with end-use quality traits. J. Food Qual. 23:391-407

Berman M, Bason ML, Ellison F, Peden G, Wrigley CW (1996) Image analysis of whole grains to screen for flour-milling yield in wheat breeding. Cereal Chemistry 73:323-327

Brachi B, Faure N, Horton M, Flahauw E, Vazquez A, Nordborg M, Bergelson J, Cuguen J, Roux F (2010) Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. PLoS Genetics 6:1

Breseghello F, Finney PL, Gaines C, Andrews L, Tanaka J, Penner G, Sorrells ME (2005) Genetic loci related to kernel quality differences between a soft and a hard wheat cultivar. Crop Science 45:1685-1695

Breseghello F, Sorrells ME (2006) Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. Genetics 172:1165-1177

Breseghello F, Sorrells ME (2007) QTL analysis of kernel size and shape in two hexaploid wheat mapping populations. Field Crops Research 101:172-179

Briarty LG, Hughes CE, Evers AD (1979) The developing endosperm of wheat - a stereological analysis. Annals of Botany 44:641-658

Brocklehurst PA (1977) Factors controlling grain weight in wheat. Nature 266:348-349

Campbell KG, Bergman CJ, Gualberto DG, Anderson JA, Giroux MJ, Hareland G, Fulcher RG, Sorrells ME, Finney PL (1999) Quantitative trait loci associated with kernel traits in a soft x hard wheat cross. Crop Science 39:1184-1195

Cane K, Sharp PJ, Eagles HA, Eastwood RF, Hollamby GJ, Kuchel H, Lu M, Martin PJ (2008) The effects on grain quality traits of a grain serpin protein and the *VPM1* segment in southern Australian wheat breeding. Australian Journal of Agricultural Research 59:883-890

Chan EKF, Rowe HC, Kliebenstein DJ (2010) Understanding the evolution of defense metabolites in *Arabidopsis thaliana* using genome-wide association mapping. Genetics 185:991-1007

Cho Y, Kang H, Lee J, Lee Y, Lim S, Gauch H, Eun M, McCouch SR (2007) Identification of quantitative trait loci in rice for yield, yield components, and agronomic traits across years and locations. Crop Science 47:2403-2417

Chojecki AJS, Gale MD, Bayliss MW (1983) Reciprocal monosomic analysis of grain size in wheat. In: Sakamoto S (Ed) Proc 6th Int Wheat Genet Symp. Maruzen, Kyoto, Japan. 1061

Cober ER, Voldeng HD, Fregeau-Reid JA (1997) Heritability of seed shape and seed size in soybean. Crop Sci 37:1767-1769

Cochrane MP (1983) Morphology of the crease region in relation to assimilate uptake and water loss during caryopsis development in barley and wheat. Aust. J. Plant Physiol. 10:473-491

Crossa J, Burgueno J, Dreisigacker S, Vargas M, Herrera-Foessel SA, Lillemo M, Singh RP, Trethowan R, Warburton M, Franco J, Reynolds M, Crouch JH, Ortiz R (2007) Association analysis of historical bread wheat germplasm using additive genetic covariance of relatives and population structure. Genetics 177:1889-1913

Dana W, Ivo W (2008) Computer image analysis of seed shape and seed color for flax cultivar description. Comput Electron Agr 61:126-135

DeSouza N (2010) High-throughput phenotyping. Nat Methods 7:36

Diao X, Furuno T, Fujita M (1999) Digital image analysis of cross-sectional tracheid shapes in Japanese softwoods using the circularity index and aspect ratio. J Wood Sci 45:98-105

Del Moral L, Garcia F, Rharrabti Y, Villegas D, Royo C (2003) Evaluation of grain yield and its components in durum wheat under Mediterranean conditions. Agronomy Journal 95:266-274

Dholakia BB, Ammiraju JSS, Singh H, Lagu MD, Roder MS, Rao VS, HS Dhaliwal, Ranjekar PK, Gupta VS (2003) Molecular marker analysis of kernel size and shape in bread wheat. Plant Breeding 122:392-395

Distelfield A, Uauy C, Fahima T, Dubcovsky J (2006) Physical map of the wheat high-protein content gene *Gpc-B1* and development of a high-throughput molecular marker. New Phytologist 169:753-763

Doehlert DC, McMullen MS, Jannink J, Panigrahi S, Gu H, Riveland NR (2004) Evaluation of oat kernel size uniformity. Crop Sci 44:1178-1186

Doehlert DC, Jannink J, McMullen MS (2006) Kernel size variation in naked oat. Crop Sci 46:1117-1123

Doganlar S, Frary A, Tanksley SD (2000) The genetic basis of seed-weight variation: Tomato as a model system. Theoretical and Applied Genetics 100:1267-1273

Drea S, Leader DJ, Arnold BC, Shaw P, Dolan L, Doonan JH (2005) Systematic spatial analysis of gene expression during wheat caryopsis development. The Plant Cell 17:2172-2185

Eathington SR, Crosbie TM, Edwards MD, Reiter RS, Bull JK (2007) Molecular markers in a commercial breeding program. Crop Sci 47:S-154

Ellis MH, Rebetzke GJ, Chandler P, Bonnett DG, Spielmeyer W, Richards RA (2004) The effect of different height reducing genes on the early growth of wheat. Funct. Plant Biol. 31:583-589

Engledow FL (1920) The inheritance of glume-length and grain-length in a wheat cross. Journal of Genetics 10:109-134

Evers AD (1971) Scanning electron microscopy of wheat starch; granule development in endosperm. Die Starke 23:157-162

Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics. Longman, Essex, England.

Fan C, Xing Y, Mao H, Lu T, Han B, Xu C, Li X, Zhang Q (2006) *GS3*, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. Theoretical and Applied Genetics 112:1164-1171

Finney PL, Andrews LC (1986) Revised microtesting for soft wheat quality evaluation. Cereal Chem 63:177-182

Fischer RA, Stockman YM (1986) Increased kernel number in Norin 10-derived dwarf wheat: evaluation of the cause. Aust. J. Plant Physiol. 13:767-784

Fischer, RA, Quail, KJ (1990) The effect of major dwarfing genes on yield potential in spring wheats. Euphytica 46:51-56

Flintham JE, Borner A, Worland AJ, Gale MD (1997) Optimizing wheat grain yield: Effects of *rht* (gibberellin-insensitive) dwarfing genes. 128:11-25

Flint-Garcia SA, Thornsberry JM, Buckler ES (2003) Structure of linkage disequilibrium in plants. Annu. Rev. Plant Biol. 54:357-374

Gegas VC, Nazari A, Griffiths S, Simmonds J, Fish L, Orford S, Sayers L, Doonan JH, Snape JW (2010) A genetic framework for grain size and shape variation in wheat. Plant Cell 22:1046-1056

Giura A, Saulescu NN (1996) Chromosomal location of genes controlling grain size in a large grained selection of wheat (*Triticum aestivum* L.). Euphytica 89:77-80

Goncharov NP, Watanabe N (2005) Physical mapping and chromosomal location of the photoperiod response gene *Ppd2* in common wheat. Breeding Science 55:81-86

Goto S, Iwata H, Shibano S, Ohya K, Suzuki A, Ogawa H (2005) Fruit shape variation in *Fraxinus mandshurica* var. *japonica* characterized using elliptic Fourier descriptors and the effect on flight duration. Ecol Res 20:733-738

Groos C, Robert N, Bervas E, Charmet G (2003) Genetic analysis of grain protein-content, grain yield and thousand-kernel weight in bread wheat. Theoretical and Applied Genetics 106:1032-1040

Guo L, Ma L, Jiang H, Zeng D, Hu J, Wu L, Gao Z, Zhang G, Qian Q (2009) Genetic analysis and fine mapping of two genes for grain shape and weight in rice. Journal of Integrative Plant Biology 51:45

Gupta, PK, Langridge, P, Mir, RR (2010) Marker-assisted wheat breeding: present status and future possibilities. Mol Breeding 26:145-161

Halloran GM (1976) Genetic analysis of hexaploid wheat, *Triticum aestivum* using intervarietal chromosome substitution lines; protein content and grain weight. Euphytica 25:65-71

Huang XQ, Cloutier S, Lyear L, Radovanovic N, Humphreys DG, Noll JS, Somers DJ, Brown PD (2006) Molecular detection of QTLs for agronomic and quality traits in a doubled haploid population derived from two Canadian wheats (*Triticum aestivum* L.) Theor Appl Genetics 113:753-766

Hedden P (2003) The genes of the Green Revolution. Trends Genet. 19:5-9

Heffner EL, Chomdej O, Williams KR, Sorrells ME (2008) Dominant male-sterile populations for association mapping and introgression of exotic wheat germplasm. Australian Journal of Agricultural Research 59:470-474

Heffner EL, Lorenz AJ, Jannink J, Sorrells ME (2010) Plant breeding with genomic selection: gain per unit time and cost. Crop Sci 50:1681-1690

Himstedt M, Fricke T, Wachendorf M (2009) Determining the contribution of legumes in legume-grass mixtures using digital image analysis. Crop Sci 49:1910-1916

Hook SCW (1984) Specific weight and wheat quality. Journal of the Science of Food and Agriculture 35:1136-1141

Horgan GW (2001) The statistical analysis of plant part appearance — a review. Comput Electron Agr 31:169-190

Houle D, Govindaraju DR, Omholt S (2010) Phenomics: the next challenge. Nat Rev Genet 11:855-866

Ibaraki Y, Kenji K (2001) Application of image analysis to plant cell suspension cultures. Comput Electron Agr 30:193-203

Igartua E, Edney M, Rossnagel BG, Spaner D, Legge WG, Scoles GJ, Eckstein PE, Penner GA, Tinker NA, Briggs KG, Falk DE, Mather DE (2000) Marker-based selection of QTL affecting grain and malt quality in two-row barley. Crop Science 40:1426-1433

Iwata H, Niikura S, Matsuura S, Takano Y, Ukai Y (1998) Evaluation of variation of root shape of Japanese radish (*Raphanus sativus* L.) based on image analysis using elliptic Fourier descriptors. Euphytica 102:143-149

Iwata H, Ukai Y (2002) SHAPE: A computer program package for quantitative evaluation of biological shapes based on elliptic Fourier descriptors. J Hered 93:384-385

Iwata H, Ebana K, Uga Y, Hayashi T, Jannink J (2010) Genome-wide association study of grain shape variation among *Oryza sativa* L. germplasms based on elliptic Fourier analysis. Molecular Breeding 25:203-215

Kato K, Miura H, Akiyama M, Kuroshima M, Sawada S (1998) RFLP mapping of the three major genes, *Vrn1*, *Q,* and *B1*, on the long arm of chromosome 5A of wheat. Euphytica 101:91-95

Kato K, Miura H, Sawada S (1999) QTL mapping of genes controlling ear emergence time and plant height on chromosome 5A of wheat. Theoretical and Applied Genetics 98:472-477

Kato K, Miura H, Sawada S (2000) Mapping QTLs controlling grain yield and its components on chromosome 5A of wheat. Theoretical and Applied Genetics 101:1114-1121

Kato K, Sonokawa R, Miura H, Sawada S (2003) Dwarfing effect associated with the threshability gene *Q* on wheat chromosome 5A. Plant Breeding 122:489-492

Keefe PD (1990) Observations concerning shape variation in wheat grains. Seed Science and Technology 18:629-640

Keyes, G (1989) *Rht1* and *Rht2* semidwarf genes effect on hybrid vigor and agronomic traits of wheat. Crop Science 29:1442

Konishi S, Izawa T, Lin SY, Ebana K, Fukuta Y, Sasaki T, Yano M (2006) An SNP caused loss of seed shattering during rice domestication. Science 312:1392-1396

Korzun V, Roder M, Worland AJ, Borner A (1997) Intrachromosomal mapping of genes for dwarfing (*Rht12*) and vernalization response (*Vrn1)* in wheat by using RFLP and microsatellite markers. Plant Breeding 116:227-232

Korzun V, Roder MS, Ganal MW, Worland AJ, Law CN (1998) Genetic analysis of the dwarfing gene (*Rht8*) in wheat. Part I. molecular mapping of *Rht8* on the short arm of chromosome 2D of bread wheat (*Triticum aestivum* L.). Theoretical and Applied Genetics 96:1104-1109

Kozak M, Madry W (2006) Note on yield component analysis. Cereal Research Communications 34:933-940

Kwack MS, Kim EN, Lee H, Kim J, Chun S, Kim KD (2005) Digital image analysis to measure lesion area of cucumber anthracnose by *Colletotrichum orbiculare*. J Gen Plant Pathol 71:418-421

Lamba, PS (1949) The relation of glume measurements to kernel shape and size in wheat. Agronomy Journal 41:167

Law CN (1967) The location of genetic factors controlling a number of quantitative characters in wheat. Genetics 56:445-461

Li C, Zhou A, Sang T (2006) Rice domestication by reducing shattering. Science 311:1936-1939

Li J, Thomson M, McCouch SR (2004) Fine mapping of a grain-weight quantitative trait locus in the pericentromeric region of rice chromosome 3. Genetics 168:2187-2195

Li JZ, Huang XQ, Heinrichs G, Ganal MW, Roder MS (2005) Analysis of QTLs for yield, yield components, and malting quality in a BC 3-DH population of spring barley. Theoretical and Applied Genetics 110:356-363

Li Q, Li L, Yang X, Warburton M, Bai G, Dai J, Li J, Yan J (2010) Relationship, evolutionary fate and function of two maize co-orthologs of rice *GW2* associated with kernel size and weight. BMC Plant Biology 10:143

Li S, Jia J, Wei X, Zhang X, Li L, Chen H, Fan Y, Sun H, Zhao X, Lei T, Xu Y, Jiang F, Wang H, Li L (2007) An intervarietal genetic map and QTL analysis for yield traits in wheat. Mol. Breed. 20:167-178

Lingle SE, Chevalier P (1985) Development of the vascular tissue of the wheat and barley caryopsis as related to the rate and duration of grain filling. Crop Sci 25:123-128

Lu C, Shen L, Tan Z, Xu Y, He P, Chen Y, Zhu L (1996) Comparative mapping of QTLs for agronomic traits of rice across environments using a doubled haploid population. Theoretical and Applied Genetics 93:1211-1217

Luo M, Dennis ES, Berger F, Peacock WJ, Chaudhury A (2005) *MINISEED3* (*MINI3*), a WRKY family gene, and *HAIKU2* (*IKU2*), a leucine-rich repeat (LRR) kinase gene, are regulators of seed size in *Arabidopsis*. Proceedings of the National Academy of Sciences 102:17531-17536

Marshall D, Ellison F, Mares D (1984) Effects of grain shape and size on milling yields in wheat. I. Theoretical analysis based on simple geometric models. Australian Journal of Agricultural Research 35:619-630

Marshall D, Mares D, Moss H, Ellison F (1986) Effects of grain shape and size on milling yields in wheat. II. Experimental studies. Australian Journal of Agricultural Research 37:331-342

Mather DE, Tinker NA, LaBerge DE, Edney M, Jones BL, Rossnagel BG, Legge WG, Briggs KG, Irvine RG, Falk DE, Kasha KJ (1997) Regions of the genome that affect grain and malt quality in a North American two-row barley cross. Crop Science 37:544-554

Matsuoka M, Ashikari M (2007) A quantitative trait locus regulating rice grain width. Nat. Genet. 39:583-584

Maystrenko O.I., Laikova L.I., Arbuzova V.S., Melnik V.M. (1998) The chromosomal location of the S1, S2, and S3 genes of induced sphaerococcoid mutations in common wheat. EWAC Newsl Proc 10th EWAC meeting, University of Tuscia, Italy:127-130

Millet E (1986) Relationships between grain weight and the size of floret cavity in the whet spike. Annals of Botany 58:417-423

Millet E, Pinthus MJ (1984) Effects of removing floral organs, light penetration and physical constraint on the development of wheat grains. Annals of Botany 53:261-269

Mohler V, Lukman R, Ortiz-Islas S, William M, Worland AJ, Beem JV, Wenzel G (2004) Genetic and physical mapping of photoperiod insensitivity gene *Ppd-B1* in common wheat. Euphytica 138:33-40

Montes JM, Melchinger AE, Reif JC (2007) Novel throughput phenotyping platforms in plant genetic studies. Trends Plant Sci 12:433-436

Motzo R, Francesco G, Giovanni P (2010) The response of rate and duration of grain filling to long-term selection for yield in Italian durum wheats. Crop & Pasture Science 61:162-169

Mou BQ, Kronstad WE (1994a) Duration and rate of grain filling in selected winter-wheat populations. 1. Inheritance. Crop Science 34:833-837

Mou BQ, Kronstad WE (1994b) Grain filling parameters and protein-content in selected winter-wheat populations. 2. Associations. Crop Science 34:838-841.

Nadaud I, Girousse C, Debiton C, Chambon C, Bouzidi MF, Martre P, Branlard G (2010) Proteomic and morphological analysis of early stages of wheat grain development. Proteomics 10:2901-2910

Novaro P, Colucci F, Venora G, D'Egidio MG (2001) Image analysis of whole grains: A noninvasive method to predict semolina yield in durum wheat. Cereal Chemistry 78:217-221

Ohsawa R, Tsutsumi T, Uehara H, Namai H, Ninomiya S (1998) Quantitative evaluation of common buckwheat (*Fagopyrum esculentum* Moench) kernel shape by elliptic Fourier descriptor. Euphytica 101:175-183

Parker GD, Chalmers KJ, Rathjen AJ, Langridge P (1999) Mapping loci associated with milling yield in wheat (*Triticum aestivum* L.). Molecular Breeding 5:561-568

Paterson AH, Damon S, Hewitt JD, Zamir D, Rabinowitch HD, Lincoln SE, Lander ES, Tanksley SD (1991) Mendelian factors underlying quantitative traits in tomato - comparison across species, generations, and environments. Genetics 127:181-197

Pearce, S, Saville, R, Vaughan, SP, Chandler, PM, Wilhelm, EP, Sparks, CA, Al-Kaff, N, Korolev, A, Boulton, MI, Phillips, AL, Hedden, P, Nicholson, P, Thomas, SG (2011) Molecular characterization of *Rht-1* dwarfing genes in hexaploid wheat. Plant Physiology 157:1820-1831

Peleman JD, Van der Voort JR (2003) Breeding by design. Trends Plant Sci 8:330-334

Peltonen-Sainio P, Kangas A, Salo Y, Jauhiainen L (2007) Grain number dominates grain weight in temperate cereal yield determination: Evidence based on 30 years of multi-location trials. Field Crops Research 100:179-188

Peng J, Ronin Y, Fahima T, Röder MS, Li Y, Nevo E, Korol A (2003) Domestication quantitative trait loci in *Triticum dicoccoides*, the progenitor of wheat. Proc. Natl. Acad. Sci. U. S. A. 100:2489-2494

Petrovic S, Worland AJ (1988) The use of reciprocal monosomic analysis to detect variation between certain chromosomes of the wheat varieties Bersee and Sava. In: Proc. 7th Int. Wheat Genetics Symposium. Cambridge, England. 629-633

Poland JA, Brown PJ, Sorrells ME, Jannink JL (2012) Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS ONE 7(2): e32253

Pritchard JK, Stephens M, Rosenberg NA, Donnelly P (2000) Association mapping in structured populations. Am. J. Hum. Genet. 67:170-181

Quarrie SA, Steed A, Calestani C, Semikhodskii A, Lebreton C, Chinoy C, Steele N, Pljevljakusi D, Waterman E, Weyen J, Schondelmaier J, Habash DZ, Farmer P, Saker L, Clarkson DT, Abugalieva A, Yessimbekova M, Turuspekov Y, Abugalieva S, Tuberosa R, Sanguineti MC, Hollington PA, Aragus R, Royo A, Dodig D (2005) A high-density genetic map of hexaploid wheat (*Triticum aestivum* L.) from the cross chinese spring x Q1 and its use to compare QTLs for grain yield across a range of environments. Theoretical and Applied Genetics 110:865-880

Radley, M (1981) The effect on wheat grain growth of the removal or ABA treatment of glumes and lemmas. J. Exp. Bot. 32:129-140

Rafalski JA (2010) Association genetics in crop improvement. Curr. Opin. Plant Biol. 13:174-180

Raju, GN, Srinivas, T (1991) Effect of husk morphology on grain development and topography in rice. Economic Botany 45:429-434

Rasmussen SK, Dahl SW, Norgard A, Hejgaard (1996) A recombinant wheat serpin with inhibitory activity. Plant Molecular Biology 30:673-677

Rebetzke, GJ, Richards, RA (2000) Gibberellic acid-sensitive dwarfing genes reduce plant height to increase kernel number and grain yield of wheat. Aust. J. Agric. Res. 51:235-245

Rebetzke GJ, Ellis MH, Bonnett D, Mickelson B, Condon AG, Richards, RA (2012) Height reduction and agronomic performance for selected gibberellin-responsive dwarfing genes in bread wheat (*Triticum aestivum* L.). Field Crops Research 126:87-96

Roberts TH, Marttila S, Rasmussen SK, Hejgaard J (2003) Differential gene expression for suicide-substrate serine proteinase inhibitors (serpins) in vegetative and grain tissues of barley. J Exp Bot 54:2251-2263

Rosenkrands I, Hejgaard J, Rasmussen SK, Bjorn SE (1994) Serpins from wheat-grain. FEBS Lett 343:75-80

Sadras VO, Egli DB (2008) Seed size variation in grain crops: Allometric relationships between rate and duration of seed growth. Crop Sci 48:408-416

Salina E, Borner A, Leonova I, Korzun V, Laikova L, Maystrenko O, Roder MS (2000) Microsatellite mapping of the induced sphaerococcoid mutation genes in *Triticum aestivum*. Theoretical and Applied Genetics 100:686-689

Sandhu D, Gill KS (2002) Gene-containing regions of wheat and the other grass genomes. Plant Physiology 128:803-811

Sayre KD, Rajaram S, Fischer RA (1997) Yield potential progress in short bread wheats in northwest Mexico. Crop Sci. 37:36-42

Scarth R, Kirby E, Law C (1985) Effects of the photoperiod gene-*Ppd1* and gene-*Ppd2* on growth and development of the shoot apex in wheat. Ann Bot 55:351-359

Schuler SF, Bacon RK, Finney PL, Gbur EE (1995) Relationship of test weight and kernel properties to milling and baking quality in soft red winter-wheat. Crop Sci. 35:949-953

Shimoji H, Tokuda G, Tanaka Y, Moshiri B, Yamasaki H (2006) A simple method for two-dimensional color analyses of plant leaves. Russ J Plant Physl+ 53:126-133

Simons KJ, Fellers JP, Trick HN, Zhang ZC, Tai YS, Gill BS, Faris JD (2006) Molecular characterization of the major wheat domestication gene *Q*. Genetics 172:547-555

Skovmand B, Reynolds MP, DeLacy IH (2001) Mining wheat germplasm collections for yield enhancing traits. Euphytica 119:25-32

Slafer GA, Savin R (1994) Source—sink relationships and grain mass at different positions within the spike in wheat. Field Crops Research 37:39-49

Snape JW, Law CN, Parker BB, Worland AJ (1985) Genetical analysis of chromosome 5A of wheat and its influence on important agronomic characters. Theoretical and Applied Genetics 71:518-526

Snape JW, Quarrie SA, Laurie DA (1996) Comparative mapping and its use for the genetic analysis of agronomic characters in wheat. Euphytica 89:27-31

Snape J, Foulkes M, Simmonds J, Leverington M, Fish L, Wang Y, Ciavarrella M (2007) Dissecting gene x environmental effects on wheat yields via QTL and physiological analysis. Euphytica 154:401-408

Somers D, Isaac P, Edwards K (2004) A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.). Theoretical and Applied Genetics 109:1105-1114

Song X, Huang W, Shi M, Zhu M, Lin H (2007) A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. Nat. Genet. 39:623-630

Sorrells ME, Fritz SE (1982) Application of a dominant male-sterile allele to the improvement of self-pollinated crops. Crop Science 22:1033-1035

Sorrells ME, LaRota M, Bermudez-Kandianis CE, Greene RA, Kantety R, Munkvold JD, Miftahudin, Mahmoud A, Ma XF, Gustafson PJ, Qi LLL, Echalier B, Gill BS, Matthews DE, Lazo GR, Chao SM, Anderson OD, Edwards H, Linkiewicz AM, Dubcovsky J, Akhunov ED, Dvorak J, Zhang DS, Nguyen HT, Peng JH, Lapitan NLV, Gonzalez-Hernandez JL, Anderson JA, Hossain K, Kalavacharla V, Kianian SF, Choi DW, Close TJ, Dilbirligi M, Gill KS, Steber C, Walker-Simmons MK, McGuire PE, Qualset CO (2003) Comparative DNA sequence analysis of wheat and rice genomes. Genome Research 13:1818-1827

Sorrells ME, Gustafson PJ, Somers D, Chao S, Benscher D, Guedira-Brown G, Huttner E, Kilian A, McGuire PE, Ross K, Tanaka J, Wenzl P, Williams K, Qualset CO (2011) Reconstruction of the Synthetic W7984 x Opata M85 wheat reference population. Genome 54:875-882

Shouche SP, Rastogi R, Bhagwat SG, Sainis JK (2001) Shape analysis of grains of Indian wheat varieties. Comput Electron Agr 33:55-76

Sourdille P, Tixier MH, Charmet G, Gay G, Cadalen T, Bernard S, Bernard M (2000) Location of genes involved in ear compactness in wheat (*Triticum aestivum*) by means of molecular markers. Mol Breed 6:247-255

Stuber CW, Polacco M, Lynn SM (1999) Synergy of empirical breeding, marker-assisted selection, and genomics to increase crop yield potential. Crop Sci 40:778-782

Sun X, Wu K, Zhao Y, Kong F, Han G, Jiang H, Huang X, Li R, Wang H, Li S (2009) QTL analysis of kernel shape and weight using recombinant inbred lines in wheat. Euphytica 165:615-624

Takano-Kai N, Jiang H, Kubo T, Sweeney M, Matsumoto T, Kanamori H, Padhukasahasram B, Bustamante C, Yoshimura A, Doi L, McCouch S (2009) Evolutionary history of *GS3*, a gene conferring grain length in rice. Genetics 182:1323-1334

Tappan JH, Wright ME, Sistler FE (1987) Error sources in a digital image analysis system. Comput Electron Agr 2:109-118

Thornsberry JM, Goodman MM, Doebley J, Kresovich S, Nielsen D, Buckler ES (2001) *Dwarf8* polymorphisms associate with variation in flowering time. Nat. Genet. 28:286-289

Tsilo TJ, Hareland GA, Simsek S, Chao S, Anderson JA (2010) Genome mapping of kernel characteristics in hard red spring wheat breeding lines. Theoretical and Applied Genetics 121:717-730

Uauy C, Brevis JC, Dubcovsky J (2006) The high grain protein content gene *gpc-B1* accelerates senescence and has pleiotropic effects on protein content in wheat RID A-4969-2008. J. Exp. Bot. 57:2785-2794

Ugalde TD, Jenner CF (1990a) Route of substrate movement into wheat endosperm. I. Carbohydrates. Aust. J. Plant Physiol. 17:693-704

Ugalde TD, Jenner CF (1990b) Rout of substrate movement into wheat endosperm. II. Amino acids. Aust. J. Plant Phyisol. 17:705-714

Varshney RK, Prasad M, Roy JK, Kumar N, Singh H, Dhaliwal HS, Balyan HS, Gupta PK (2000) Identification of eight chromosomes and a microsatellite marker on 1AS associated with QTL for grain weight in bread wheat. Theoretical and Applied Genetics 100:1290-1294

Wang R, Hai L, Zhang X, You G, Yan C, Xiao S (2009) QTL mapping for grain filling rate and yield-related traits in RILs of the Chinese winter wheat population Heshangmai x Yu8679. Theoretical and Applied Genetics 118:313-325

Waters BM, Uauy C, Dubcovsky J, Grusak MA (2009) Wheat (*Triticum aestivum*) NAM proteins regulate the translocation of iron, zinc, and nitrogen compounds from vegetative tissues to grain RID A-4969-2008. J. Exp. Bot. 60:4263-4274

White RJ, Prentice HC, Verwijst T (1988) Automated image acquisition and morphometric description. Can J Bot 66:450-459

Wiersma JJ, Busch RH, Fulcher GG, Hareland GA (2001) Recurrent selection for kernel weight in spring wheat. Crop Science 41:999-1005

Willcox G (2004) Measuring grain size and identifying near eastern cereal domestication: Evidence from the Euphrates valley. Journal of Archaeological Science 31:145-150

Wilson WA, Harrington SE, Woodman WL, Lee M, Sorrels ME, McCouch SR (1999) Inferences on the genome structure of progenitor maize through comparative analysis of rice, maize and the domesticated panicoids. Genetics 153:453-473

Xie X, Jin F, Song M, Suh J, Hwang H, Kim Y, McCouch S, Ahn S (2008) Fine mapping of a yield-enhancing QTL cluster associated with transgressive variation in an *Oryza sativa* x *O. rufipogon* cross. Theoretical and Applied Genetics 116:613-622

Yamazaki WT, Briggle LW (1969) Components of test weight in soft wheat. Crop Science 9:457-459

Yan C, Yan S, Yang Y, Zeng X, Fang Y, Zeng S, Tian C, Sun Y, Tang S, Gu M (2009) Development of gene-tagged markers for quantitative trait loci underlying rice yield components. Euphytica 169:215-226

Yu J, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, Kresovich S, Buckler ES (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat. Genet. 38:203-208

Yu J, Holland JB, McMullen MD, Buckler ES (2008) Genetic design and statistical power of nested association mapping in maize. Genetics 178:539-551

Zhang LY, Liu DC, Guo XL, Yang WL, Sun JZ, Wang DW, Zhang A (2010) Genomic distribution of quantitative trait loci for yield and yield-related traits in common wheat. Journal of Integrative Plant Biology 62:996-1007

CHAPTER TWO

Comparison of digital image analysis using elliptic Fourier descriptors and major dimensions to phenotype seed shape in hexaploid wheat (*Triticum aestivum* L.)

Abstract:

Digital image analysis (DIA) is widely used for describing plant organ shape. However, the various types of shape descriptors that can be generated using DIA may identify different loci in genetic analyses. The purpose of this study was to evaluate two different DIA approaches to quantifying wheat seed shape for exploring trait correlations and quantitative trait loci (QTL) mapping. Phenotypic data were produced using the software programs ImageJ (National Institutes of Health, USA, http://rsbweb.nih.gov/ij/) and SHAPE (Hiroyoshi Iwata, http://lbm.ab.a.u-tokyo.ac.jp/~iwata/shape/). ImageJ generates measures of length, width, perimeter, and area that can be used to describe dimensions of objects, whereas SHAPE generates elliptic Fourier descriptors (EFDs) to capture shape variation such as roughness, asymmetric skewing, or other two-dimensional aspects not encompassed by axes or distinctions in overall object area. There were significant differences in the results of the QTL analysis depending on the DIA software used. The use of EFDs to characterize horizontal measures of seed shape in wheat identified more QTL with higher LOD scores than length to width ratio. Additionally, the entire three dimensional shape of the seed described using two images in different orientations was shown to identify seed shape QTL that co-located with flour yield and would go undetected based solely on a two dimensional image of the seed. Both methods identified QTL for length, width, thickness, and vertical perimeter that were co-localized with QTL for flour yield.

I. Introduction

In recent years advances in phenotypic measurements have fallen behind progress in high-throughput genotyping. In order to take full advantage of low-cost genotyping resources plant breeders are aggressively pursuing more accurate and efficient phenotyping methods (Montes *et al.* 2007; DeSouza 2010; Houle *et al.* 2010). Refinement of photometric approaches will provide more concise, potentially cheaper, phenotypic information and better elucidate the role of individual genetic components of complex traits. However, different approaches to converting raw images of plant organs into quantitative data could influence the results of genetic analyses.

Digital image analysis (DIA) is the process of converting digital images of individual objects, such as plant organs, into quantitative measurements (Diao *et al.* 1999; Horgan 2001; Ibaraki and Kenji 2001; Shouche *et al.* 2001; Kwack *et al.* 2005; Dana and Ivo 2008; Himstedt *et al.* 2009). DIA can increase the speed of phenotyping, allowing the rapid generation of large quantitative data sets. DIA methods that convert photographs of plant organs into quantitative data based on measures of axes or pixel counts have been used by numerous research groups (Cober *et al.* 1997; Campbell *et al.* 1999; Doehlert *et al.* 2004; Kwack *et al.* 2005; Breseghello and Sorrells 2006; Shimoji *et al.* 2006; Breseghello and Sorrells 2007; Dana and Ivo 2008). Despite these advantages, there is little information regarding optimization of the photographic process and the impact of image analysis methodology on experimental results.

There are two common methods of converting images to quantitative data; the measurement of dimensions of an object and mathematical descriptors of shape. Measurements of object dimensions using major axes or pixel counts can be obtained using computer programs such as ImageJ (National Institutes of Health, USA, http://rsbweb.nih.gov/ij/) or Photoshop (Adobe Systems Incorporated, www.photoshop.com). Mathematical descriptors capture aspects of shape that are not readily described by measuring dimensions, such as asymmetry, undulation of leaf margins, or surface texture. White *et al.* (1988) suggested the use of elliptic Fourier descriptors (EFDs) to capture these

characteristics and provide robust quantitative measures of plant organ shape. EFDs are generated by superimposing the outline of a shape onto a coordinate plane then converting the outline into a numeric description that can be subjected to principal component analysis (PCA). Individual PCA scores can then be used directly as phenotypic data for genetic analyses. Thus, EFDs have found use in several studies of plant organ morphology owing to their ability to quantify aspects of shape that were formerly limited to categorical description or rough estimation (Iwata *et al.* 1998; Ohsawa *et al.* 1998; Goto *et al.* 2005). For example, the program SHAPE (Hiroshi Iwata, http://lbm.ab.a.u-tokyo.ac.jp/~iwata/shape) readily generates principal component scores (PCAs) from EFDs (Iwata and Ukai 2002).

Studies of kernel morphology can exploit increasingly refined phenotyping methods. In wheat, such measurements of seeds may relate to traits such as yield or milling quality, which are economically important but costly to evaluate. Based on geometric models it has been suggested that wheat seed shape might influence flour yield, since spherical seeds will have the highest possible endosperm to bran ratio (Marshall *et al.* 1984; Novaro *et al.* 2001). However, this relationship has been debated in the literature (Marshall *et al.* 1986) and it is complicated by the crease that runs along the dorsal side of the seed.

The use of DIA provides an opportunity to re-evaluate both the genetic and phenotypic components of seed shape. In this study, several methods of assessing wheat seed shape were evaluated. These included (1) measures of axes and pixel counts using ImageJ, (2) EFD measures of shape using SHAPE, and (3) derived measures of shape that incorporate multiple measures from (1), (2), and seed weight. Subsequent QTL analyses were then performed using each of these measures. QTL for seed shape characteristics, flour yield, and TKW were compared. Additionally, phenotypic scores of seed shape were tested for correlation with each other and for relationship to the quality trait flour yield (FLYLD).

II. Materials & Methods

Plant Materials

Seed from the Cayuga x Caledonia doubled haploid (CxC) mapping population consisting of 208 doubled-haploid lines was used for phenotyping. Both parents are winter wheat cultivars with good agronomic qualities, and are adapted to the Northeastern U.S. Details of the development of the CxC population can be found in Munkvold *et al.* 2009, and seed is available through the U.S. WheatCAP project (http://maswheat.ucdavis.edu/). A subset of 161 lines from three environments (Helfer, McGowan, and Ketola fields) grown in 2005 was included for seed morphology measures based on quality of available seed and genotypic data. Flour yield measurements were taken on lines grown in Snyder field during the 2005, 2006, and 2008 growing seasons. Each location consisted of two replicates of single one-meter rows grown in a randomized complete block design (Munkvold *et al.* 2009). Rows were individually harvested by hand and threshed using a mechanical belt thresher.

Phenotyping

Photography and pre-processing: Seed photography was adapted from the work of Breseghello and Sorrells (2007). For each entry, 25 sound, intact kernels were selected for photographing. Undamaged, non-shriveled kernels which excluded the occasionally extremely large or extremely small kernels seen in some threshed samples were included as representative of each line. These kernels were laid out on black clay in a 5x5 grid spaced ~1cm apart and photographed vertically and horizontally. The two photographs included both a view of the kernel with its crease side down, referred to as the horizontal image or 'H image', and a view of the kernel positioned with the embryo end embedded in the clay, referred to as the vertical image or 'V image' (See Figure 2.1, 2.2). The seed packet for each entry was included in the photograph with genotype, environment, and plot number visible. A size standard measuring $9cm^2$ was placed below the kernels. Photos were taken using a digital

camera, and transferred to a portable notebook computer as standard Joint Photographic Experts Group (JPEG) files.

Pre-processing each photo involved renaming the files with appropriate genotype, environment, and plot identifiers, then cropping the photo to include only kernels and size standard. Cropping and renaming images was performed using the image editing software program GIMP (GNU Image Manipulation Program, www.gimp.org). All images had their contrast and brightness enhanced to reduce edge detection errors from shadowing. Contrast and brightness adjustments were performed as batch conversions using the software IrfanView (Irfan Skiljan, www.irfanview.com). These cropped and contrast-enhanced photos constituted 'raw images' that were converted to seed measurements using the image analysis programs (ImageJ and SHAPE).

Phenotyping shape using axes and pixel counts: ImageJ performs object counts and two-dimensional measurements of each object directly from JPEG files. First, JPEG images were opened in ImageJ as an image stack and converted to 8-bit, black/white binary images using the 'Make Binary' command. To derive quantitative measures from binary images, a global scale was set using the size standard included with each photograph so that ImageJ could calculate actual distance based on pixel measurements. The 'Count Object' command was used to return values for four primary measures including major axis, minor axis, area, and perimeter of each seed (Figure 2.3). For H images the major axis corresponded to seed length and minor axis corresponded to seed width. For vertical images (kernel photographed on end, Figure 2.2), conversion of binary images to quantitative measures was repeated, with the major axis corresponding to seed width and minor axis corresponding to seed thickness. Assignment of axes from the separate images to the appropriate dimensions was checked by comparison of the returned values, where width of individual seeds remained consistent between both sets of images (Figure 2.3). ImageJ output for the measures of image sets was exported to a

spreadsheet where values for seed images with poor outlines were removed based on visual inspection, and remaining measures were averaged to return phenotypic values for each genotype in each environment.

Phenotyping shape using EFDs: SHAPE converts the shapes of objects into data suitable for genetic analysis by recording elliptical Fourier descriptors (EFDs) for the objects, performing principle component analysis (PCA) using these EFD values, and returning PCA scores for each object. PCA scores summarize the large number of EFD coefficients generated for each shape and reduce them to quantitative values able to be used directly for genetic analysis by performing a PCA using the variance-covariance matrix of the coefficients. Further details of the individual steps can be found at the SHAPE website, http://lbm.ab.a.u-tokyo.ac.jp/~iwata/shapehtml. Digital photographs were converted from JPEG to binary images in bitmap file format (BMP) using the program IrfanView (www.irfanview.com). Following file conversion, horizontal and vertical image sets were processed sequentially through SHAPE, with poor outlines removed based on visual inspection prior to PCA. Individual sets of EFD files were combined into a single EFD file including all genotypes and environments used for this study prior to PCA using SHAPE, and final PCA scores for shape descriptors were exported as text files. The PCA scores for individual seeds were averaged to provide a representative PCA value for each genotype in each environment. Visual representations of the PCA scores returned from both the H image set and V image set of CxC are shown in Figure 2.4 and Figure 2.5. These measures are referred to as HPC1 through HPC5 and VPC1 through VPC5, to denote horizontal principle components and vertical principle components, respectively.

Interpretation of HPC values (see Figure 2.4) in this study are as follows: HPC1 appeared to affect the overall widening along the length of the seed and the length to width aspect of a seed, HPC2 appeared to affect widening of the seed at either the end of the seed having the embryo or the opposite

side, HPC3 appeared to affect the widening of the seed at both ends (making it more or less barrel-shaped), HPC4 appeared to affect asymmetrical widening on either side of the embryo, and HPC5 appeared to affect widening of the seed at the end opposite the embryo. Interpretation of the VPC values (see Figure 2.5) in this study are as follows: VPC1 appeared to represent the left/right orientation of the seed during photographing, VPC2 appeared to represent the depth of the kernel crease, VPC3 appeared to represent how 'V-shaped' a kernel was as well as crease depth, VPC4 appeared to represent flaring of each half of the seed adjacent to the crease, and VPC5 appeared to represent asymmetric skewing to the left or right of the crease. The first three PCs which were believed to describe the most shape variation as shown in Figures 2.4 and 2.5 were chosen for comparison. For correlation and QTL comparison, only HPC1 – HPC3 and VPC2 – VPC4 were used since each subsequent PC explains a smaller portion of the variation. VPC1 was omitted because it most likely describes the slight difference in positioning of kernels for photographing rather than true shape variation; this issue relates to the normalization of shapes and is described in the 'Discussion' section.

Phenotyping shape using derived measures: Previously, thousand kernel weight (TKW) and factor form density (FFD) have been used to describe size and shape of seeds (Giura *et al* 1996; Guo *et al* 2009). Thousand kernel weight was measured by weighing all seed from a sample, dividing by total number of seeds, and multiplying by 1,000. Factor form density was calculated as:

$$\text{FFD} = \text{grain weight} \div (\text{grain length} \times \text{grain width})$$

Several derived geometric measures of seed shape were computed from measurements recorded by ImageJ and SHAPE (Table 2.1). The volume of seeds was approximated as VOL_{XYZ} using the formula for volume of an ellipsoid (Eric W. Weisstein, Ellipsoid, From MathWorld:

http://mathworld.wolfram.com/Ellipsoid.html) based on x, y, and z axes corresponding to seed width, seed length, and seed thickness measures (respectively) from ImageJ output:

$$\mathrm{VOL_{XYZ}} = \left(\frac{4}{3}\right)\pi xyz$$

The deviation of an individual seed from an optimal ellipse was calculated based the major and minor axes of either the horizontal image (PDEVH) or the vertical image (PDEVV) (Eric W. Weisstein, Ellipsoid, From MathWorld: http://mathworld.wolfram.com/Ellipsoid.html).

$$p \approx \pi\left[3(a+b) - \sqrt{(3a+b)(a+3b)}\right]$$

where,

p = optimal perimeter value of ellipse

a = major axis of seed taken from the image

b = minor axis of seed taken from the image

From p, the 'optimal' perimeter value, the actual perimeter measurement from ImageJ was subtracted and divided by the actual perimeter to normalize for differences in overall length of perimeter (seed size). The absolute value of this was then taken to return a positive value that could quantify how closely the perimeter matched one of an ideal, smooth elliptical seed. For example, a seed with rough surface texture returned a higher value than a seed with a smooth surface.

$$\mathrm{PDEVH} = |(p - \mathrm{HPERIM})/\mathrm{HPERIM}|$$

$$\mathrm{PDEVV} = |(p - \mathrm{VPERIM})/\mathrm{VPERIM}|$$

Composite values incorporating both the derived volumetric based on primary seed axes (VOL_{XYZ}) and the fine-scale shape descriptors of perimeter deviation from ImageJ (PDEVH, PDEVV) or PCA scores returned from SHAPE were calculated. For describing a composite phenotype for kernel shape using only ImageJ data (COMPS1), phenotypic scores were calculated as:

$$\mathrm{COMPS1} = (\mathrm{VOL_{XYZ}} \div \mathrm{PDEVH}) \div \mathrm{PDEVV}$$

This COMPS1 value represents an approximated volume of the kernel as calculated using major and minor axes from both horizontal and vertical images, modified by how much the perimeter values from each profile deviated from a smooth, ideal elliptical outline. Therefore, if a kernel had a profile that was less elliptical in the horizontal orientation or vertical orientation it would be further from the idealized 'optimal' ellipsoid described by VOL_{xyz}, and would have a lower COMPS1 value. For describing a composite phenotype for kernel shape using both ImageJ data and EFD descriptors from SHAPE, a similar second shape composite denoted COMPS2 was calculated as:

$$\mathrm{COMPS2} = \mathrm{VOL_{XYZ}} \times \mathrm{HPC}a \times \mathrm{VPC}b$$

Flour Yield

Flour yield (FLYLD) data was produced by the USDA Soft Wheat Quality Lab in Wooster, OH using the Quadrumat mill as described previously (Finney and Andrews 1986; Andrews 2002).

Genotyping and Linkage Map Construction

A subset of 161 lines of the CxC population was previously genotyped at 320 loci, including data from 191 simple sequence repeat (SSR), 15 restriction fragment length polymorphism (RFLP), 31 target region amplification polymorphism (TRAP), and 72 amplified fragment length polymorphism (AFLP), 8 expressed sequence tag-SSR (EST-SSR), and 3 sequence tagged site (STS) markers. Details of the genotyping and map construction were reported by Munkvold *et al. (*2009). QTL and associated markers from these linkage groups were assigned to wheat chromosomes based on information in the GrainGenes database (Agricultural Research Service, US Department of Agriculture, www.graingenes.org) and their location on a wheat consensus map (Somers *et al.* 2004).

Data Analyses

All phenotypic measures of seed shape from ImageJ were tested for normal distribution, outliers were removed, and ANOVA was performed using JMP8.0 software. Milling quality data were used directly as individual scores for each environment. Pearson correlation coefficients between phenotypes were calculated using average line values for each trait from each environment and tested for significance by comparing observed r-values to critical r-values for a two-tailed test at significance levels of 0.05 and 0.01 for df = 159.

QTL analysis was performed using QTL Cartographer version 2.5 (North Carolina State University, www.statgen.ncsu.edu/qtlcart/). Traits were analyzed first by single marker regression analysis using all markers to test for linkage groups containing at least one significant locus for any trait

at the significance level of 0.05. From this, a reduced version of the map, including only linkage groups containing at least one locus significant for any trait from the single marker regression analysis was analyzed using composite interval mapping (CIM). Significance thresholds were set using permutation testing based on 1,000 permutations with significance threshold of $p<0.01$ prior to CIM QTL analysis. QTL Cartographer parameters were set to CIM Model 6, with 10 control markers, window size set to 10cM, a walk speed of 2.0cM, and analysis run using a Forward Regression Model. Each individual QTL region was defined as the region over which the significance line for a trait exceeded the threshold determined by permutation testing. Bordering makers were selected as the two markers closest to either edge of the significance line for a trait once it dropped below the threshold determined by permutation testing for that trait. In cases where one of the bordering markers was located close to the highest LOD peak for the QTL, the marker was also designated as the nearest marker.

III. Results

Phenotypic Correlations

Correlation coefficients between measures of seed shape recorded by ImageJ and SHAPE were used to compare the two methods (Table 2.2). For different shape characteristics, these values ranged from no correlation to significant correlations of up to 0.78 (VPC4_mcg and VPC2_mcg). The most highly correlated, different kernel morphology traits across environments were seed width and thickness (average correlation of 0.42 across 9 comparisons), HPC1 and seed length (average correlation of -0.63 across 9 comparisons), HPC1 and seed width (average correlation of 0.36 across 9 comparisons), HPC1 and HPC3 (average correlation of -0.28 across 9 comparisons), and HPC3 and length (average correlation of 0.39 across 9 comparisons). Higher correlations, with significant values as high as 0.83 and 0.84, were seen for the single traits including length and HPC1 (respectively) across different environments.

Correlation coefficients between derived measures of seed shape are listed in Table 2.3. The highest significant correlation was -0.84 between COMPS1_mcg and PDEVV_mcg. The most highly correlated derived measures across different environments were FFD and ASPECT (average correlation of -0.31 across 9 comparisons), and COMPS2 and ASPECT (average correlation of -0.16 across 9 comparisons).

The highest correlations between flour yield and direct measures of seed shape were 0.24 for HPC2_mcg and VPC4_helf (Table 2.2). All direct measures of seed shape had at least one significant correlation with flour yield, though these were generally low. The highest correlation between flour yield and a derived measure of seed shape was 0.24 for PDEVH_helf. All derived measures of seed shape had at least one significant correlation to flour yield with the exception of COMPS1, though these were also generally low. Phenotypic data on lines is listed in Supplementary Table 2.2 in Online Resources.

Quantitative Trait Loci Detected Using ImageJ

One-way ANOVA revealed that lines were significantly different for all phenotypic traits measured (p-values <0.001). From the three environments sampled, 26 QTL were detected based on ImageJ measurements of seed shape (Table 2.4, Figure 2.6). The proportion of variation explained by each of these ranged from 7 to 21% with an average of 11%. Six QTL were detected on chromosome on 2A, three on 2B, one on 3A, two on 3B, three on 3D, one on 4B, three on 4D, two on 5D, four on 6D, and one on 7D. LOD scores for these QTL ranged from 3.7 to 18, with an average of 6.17. The confidence intervals for these QTL ranged from 1.8cM – 35.4cM, with an average of 18.8cM (Table 2.5).

Quantitative Trait Loci Detected Using SHAPE

Based on the one-way ANOVA lines were significantly different for all phenotypic traits (p-values <0.001) except for VPC1 and VPC2. Using the PCA scores derived from EFD measures of seeds, 13 QTL were detected (Table 2.4, Figure 2.6). Significant QTL included one on chromosome 1A, two on 2B, one on 2D, one on 3A, two on 3B, two on 3D, one on 4D, one on 5D, and two on 6D. Values of r^2 ranged from 7% to 20%, with an average of 10%. LOD scores for these QTL ranged from 4.3 to 16, with an average of 8.15. Confidence intervals ranged from 8cM – 49cM, with an average of 25cM (Table 2.5).

QTL Detected using Derived Measures

The one-way ANOVA showed that lines were significantly different for all phenotypic traits (p-values <0.001). Using derived measures of seed shape, 14 QTL were detected (Table 2.4, Figure 2.6). Significant QTL detected included one on 2B, one on 3A, one on 3B, four on 3D, four on 4D, two on 6D, and one on 7D. Values of r^2 ranged from 5% to 18%, with an average of 11%. LOD scores ranged from 3.1 to 11.5 with an average of 6.71. Confidence intervals ranged from 2cM to 29cM, with an average of 20.4cM (Table 2.5).

Comparison of QTL Detected

The different phenotypic measures of the two DIA methods identified many of the same QTL that affected specific aspects of seed shape but could be described in multiple ways. For example, seed length is described by direct length measured using ImageJ, but is also captured in the HPC1 measurement from SHAPE. Many of the QTL detected by the two DIA programs were in similar chromosome locations (Table 2.6). In the case of some QTL, such as those detected on chromosome 7D, orientation of the seed during photography was important to detect QTL which co-localized for direct measures (VPERIM), derived measures (TKW), and quality characteristics (FLYLD) (Figure 2.7).

Derived measures of seed morphology closely mirrored the locations of their component measures. In the QTL cluster for direct measures on chromosome 2B, the derived measure PDEVH was also located there. The LOD score (LOD = 8) for PDEVH at that chromosome location was comparable to the highest significant LOD score from a direct measurement (HPC1, LOD = 7), although the confidence interval for the derived measure was larger. In the 3A cluster, the derived measure, ASPECT, co-localized with the direct measures of length and HPC1, with an intermediate LOD between length and HPC1. On chromosome 3B, ASPECT was co-located with direct measures affecting both length and width, but again had a lower LOD score than HPC1. In the QTL cluster on 3D, QTL for the derived measures of COMPS1, PDEVH, ASPECT, and VOL_{XYZ} were located. The derived measures had lower LOD scores (6.1 for derived compared to 10.4 for direct) but smaller confidence interval values (16cM for derived compared to 23cM for direct). In the region of 4D containing the *rht-D1* locus, COMPS2, FFD, ASPECT, and TKW QTL co-localized. Chromosome 6D had significant QTL for the derived measure ASPECT. Significant QTL for TKW was detected on 7D and on 2B. Notably, the QTL which co-localized with flour yield on 7D included only shape characteristics captured from the vertical images (VPERIM) and TKW. No QTL that were detected using horizontal images of seeds co-localized with the flour yield QTL on 7D. (Table 2.4, Figure 2.6). QTL which were detected are listed based on sequential chromosome order in Supplementary Table 2.2 in Online Resources.

IV. Discussion

Dissecting Seed Morphology Relationships through Correlations and QTL Analysis

There were multiple low but significant correlations between measurements of major dimensions and measurements of seed shape captured by EFDs (Table 2.2). These low correlations between the major dimensions of seeds and EFDs indicate that different aspects of seed morphology were captured by each phenotyping method and likely could be selected independently. Comparison of

the two DIA methods revealed that direct measures from ImageJ were significantly correlated with flour yield in 4 of 27 (14.8%) cases (Table 2.2), whereas SHAPE measures were correlated with flour yield in 13 of 51 (35%) cases. For comparison, TKW was significantly correlated with flour yield in 2 of 9 (22%) cases. The higher percentage of statistically significant correlations observed between seed EFD measures and flour yield suggests that they were better for relating wheat kernel morphology to flour yield. The significant correlation between HPC2 from the McGowan location and the PDEVH measure from the Helfer location with flour yield (both 0.24) provide evidence that variation in shape represented by the curvature of the seed at either end (i.e. widening at a point other than the central portion of the seed,) may impact flour yield, as was suggested by Marshall *et al.* (1984). The measures VPC3, VPC4, and PDEVV had significant correlations with flour yield that may relate to the depth of a seed's crease, for which variation is not described when using only direct measures of major axes of seed dimensions. Because EFDs were more highly correlated with FLYLD than measurements of major or minor axes, EFDs would be preferred for phenotyping if kernel shape were used in selection to increase flour yield. Furthermore, the correlations between EFDs and FLYLD suggest that they are able to relate the uniformity and smoothness of the kernel to flour yield because roughness or shriveling would be expected to reduce the ratio of internal volume to surface area of the kernel. Use of EFDs recorded from kernels imaged on end (vertical images in this study) also can characterize variation in the depth or angle of a wheat seed's crease which will impact the volume to surface area relationship of a seed.

QTL affecting seed size can be found across all chromosomes of wheat, with varying degrees of effect seen for individual QTL (Campbell *et al.* 1999; Dholakia *et al.* 2003; Breseghello *et al.* 2005; Quarrie *et al.* 2005; Sun *et al.* 2009; Gegas *et al.* 2010; Tsilo *et al.* 2010). A recent meta-QTL study has compiled the results of many of these and identified regions on 1A, 1B, 2A, 2D, 3B, 4A, 4B, 4D and 5A that are frequently cited as influencing seed morphology (Zhang *et al.*, 2010). In this study, multiple QTL

related to seed morphology were co-localized with each other and with QTL for flour yield (Table 2.6). Notably, many of the QTL detected correspond to the meta-QTL Zhang *et al.* reported on chromosomes 2A, 3B, and 4D. The QTL on 4D all co-localize with the reduced-height gene *Rht-D1*, between markers *Xbarc0217* and *Xbarc1118*. This shows the pleiotropic effect of *Rht-D1* on kernel morphology. The QTL listed in Table 2.6 can be categorized as those affecting 1) individual dimensions of the seed and flour yield, 2) multiple dimensions of the seed (meaning a single QTL that affects more than one dimension of the seed, such as length and width simultaneously), and 3) individual dimensions of the seed but not flour yield. Three QTL regions for individual dimensions defining seed shape also overlapped with QTL for flour yield and could be detected using either major dimensions of the seed or EFDs. On 2B QTL affected both width and flour yield, but there was no significant QTL for seed length detected. Notably, this cluster on 2B is distinct from the predicted region for the photoperiod sensitivity gene *Ppd-B1*, which would be near marker *Xgwm429* based on comparison to previous reports on *Ppd-B1* and linkage maps published in the GrainGenes database (Mohler *et al.* 2004, Agricultural Research Service, US Department of Agriculture, www.graingenes.org). On 2D, QTL for VPERIM and HPC1 did occur near marker *Xgwm429*, though again, these were distinct from the QTL affecting width and flour yield. On chromosome 3A a second group included QTL for seed length, HPC1, and flour yield suggesting that seed length affected flour yield. On chromosome 7D, a QTL cluster for flour yield and vertical perimeter (VPERIM) was detected by ImageJ suggesting that flour yield was related to a QTL for crease depth. These three QTL which affect both quality characteristics and specific dimensions of the seed may represent unique targets for future marker assisted selection projects.

By comparing results from both DIA programs, QTL were identified affecting shape of the seed in all pairwise combinations of major axes. Several QTL clusters on chromosomes 2A, 3B, and 6D may be harboring genes affecting development of the wheat kernel in multiple dimensions. The 2A region affected both the width and thickness of seeds. The 3B QTLs affected thickness and length of seeds. On

6D, the presence of length and width QTL suggested that these traits were more important than thickness. No QTL were detected that affected all three dimensions of seed shape. These observations imply that the overall volume of a seed (as composed of the three major dimensions) is impacted by genes affecting either one or two dimensions of the seed. This suggestion has been supported in other crops such as rice, where studies of seed shape have recently discovered QTL affecting multiple dimensions of the seed (length and width) simultaneously (Shao *et al.* 2012; Qiu *et al.* 2012). Since derived measures of seed size such as TKW are metrics in which changes in any one of the three dimensions of a seed may generate a change in overall volume, genetic studies which use TKW may identify different QTL underlying 'seed size' since they may detect different loci affecting any one of the individual component dimensions. Kernel characteristics have a moderate to high heritability, with size (as TKW) ranging from broad-sense heritability of 0.58 to 0.90 and shape parameters (length, width) ranging from 0.55 to 0.95 (Barnard *et al.* 2002; Sun *et al.* 2009; Wang *et al.* 2009; Gegas *et al.* 2010; Tsilo *et al.* 2010). In general, the trend seems to be that the TKW of cultivars is more heritable than shape parameters and length of kernels is more highly heritable than width (Sun *et al.* 2009). From a physiological standpoint, the observation of differential heritability fits evidence supporting sequential development of yield components (Kozak and Madry 2006). Length of a seed is set earlier in the developmental process whereas the width of a seed has more time to be influenced by environmental conditions during the seed filling period (Sadras and Egli 2008). That TKW is generally more highly heritable could be the result of separate temporal environmental effects on specific individual yield components being moderated or balanced by the other component traits which contribute to TKW. Based on the results of this study showing influence of specific QTL on multiple seed dimensions, selection work focused on preferentially retaining alleles which simultaneously increase length and width (6D), and those which increase length and thickness (3B) could improve the overall sink potential

of cultivars and limit the risk associated with environmental variability occurring during the grain fill period.

The QTL observed in this study suggest a model for development of seed shape from multiple genes that affect one or two dimensions of the seed. While there was a QTL for VOL_{xyz} detected on 3D, length was the only component trait that co-located with the VOL_{xyz} QTL near marker *Xcfd70*. Similar comparisons between TKW and seed dimensions show that QTL with large effects on individual dimensions may underlie TKW QTL. This supports the work of Gegas *et al.* (2010), which found seed shape when measured in two dimensions (length and width) to be independent of seed size as measured using TKW. Selection for 'seed size' using derived measures adopted in the past, such as TKW, without more refined shape description would encompass all three planes and reduce the specificity of selection for sites increasing multiple dimensions of the seed simultaneously. Selection of multiple QTL affecting a single major dimension (axis) of the seed but with variable pleiotropic effects on other traits could facilitate selection of genotypes that positively affect two dimensions of the seed simultaneously or impact quality traits such as flour yield.

The EFD descriptor HPC1 was able to detect more QTL with higher LOD scores than the derived measure ASPECT, or length to width ratio, which has been used for characterizing seed shape (Guo *et al.* 2009). In 4 out of 5 cases, HPC1 co-localized with ASPECT but returned higher LOD scores (Table 2.4). Furthermore, HPC1 was able to detect two more QTL than ASPECT (on 2A and 2B). This implies that HPC1 is a better phenotypic measure than ASPECT for detecting the genes that influence the horizontal shape of wheat seeds. Direct measures of seed shape detected more QTL, generally with higher r^2 and LOD values, than use of derived measures evaluated in this study. As such, the use of direct measures of shape including both major axes (ImageJ) and EFD descriptors (SHAPE) is recommended over derived measures for future studies examining kernel morphology.

In this study, we characterized relationships between seed dimensions and flour yield by measuring all three dimensions of the seed. Measurement of thickness of seeds via vertical images identified QTL that would otherwise go undetected using only horizontal images (Figure 2.7). A QTL region on chromosome 7D significantly influenced seed vertical perimeter but not length or width and also co-located with a flour yield QTL. Furthermore, phenotypic selection of individual dimensions of the seed may not improve flour yield, because QTL affecting seed dimensions did not always co-localize with those for flour yield.

Application of DIA Using ImageJ or SHAPE

A study using DIA must consider image quality, and in the case of SHAPE, orientation of objects and data collection. Some of these limitations have been identified in the previous literature (Tappan *et al.* 1987). The most frequent cause of errors in the phenotypic data was poor image quality that prevented the software from accurately detecting the boundaries of the seed. Uniform lighting provided by a photography stage with multiple light sources, matte background material, and high quality seed samples are critical to preventing data loss post-imaging. In addition, SHAPE detects subtle variations in an object's form and normalizes for the effects of seed orientation. SHAPE can do this based on the longest axis or the 'first harmonic' measure. While either of these work well for oblong objects with a high aspect ratio (such as wheat seeds photographed in a horizontal orientation, elongated radishes, snap beans, etc.) or objects that have radial symmetry (cross-sections of spherical melons, round seeds, etc.), proper normalization can be difficult for asymmetric objects or those which have variable bilateral symmetry (including wheat seeds photographed in a vertical orientation, many flowers, vertically constricted pumpkins, etc.). Other studies of plant organ shape using EFDs often did not address the problem of normalization, because they only examined rounded or oblong materials (Iwata *et al.* 1998; Goto *et al.* 2005; Iwata *et al.* 2010). The second consideration unique to SHAPE is

that it performs a PCA to return scores describing shape prior to subsequent analysis. Because of this, all objects to be analyzed in subsequent genetic studies must be processed through the PCA in a single step. Data produced from independent PCAs run in SHAPE cannot be grouped for subsequent genetic analysis because the individual PCA scores for each object are determined by the EFD values of the entire group. PCA scores for lines in this study cannot be directly compared to PCA scores from a study using different materials because PCA scores from a different set of images processed via SHAPE may detect a different aspect of shape depending on the variation in the population. This is a result of prioritization of principle components by SHAPE, where principle components are assigned a number in decreasing order of how much variation they explain. For example, the HPC5 (Figure 2.4) represents apical tapering in this study, whereas HPC5 could represent a different aspect of the seed, such as length to width ratio, in a different mapping population. It is important to consider both of these peculiarities of SHAPE prior to DIA to derive the most accurate information from digital images. Details of dealing with these considerations are further described in Online Resources.

The Role of Seed Size and Shape in Wheat Breeding

Knowledge of which dimensions of the seed impact quality and agronomic characteristics is important for directed manipulation of seed morphology to improve yield or quality improvements in wheat. This can be accomplished by characterization of the three dimensional shape of wheat seeds and QTL analysis to inform the selection process, whether it is based on phenotypic or genotypic methods. Empirical studies in target populations should be used to prioritize phenotypes and markers to be used in building selection indices on a larger scale. Low correlations between components of a seed's three dimensional shape indicate that these components can be manipulated independently. Marker sets can be identified flanking QTL that affect multiple dimensions of the seed or influence flour yield and can be used to aid selection in early generations or in non-target growing environments

(greenhouses, off-season nurseries). Most importantly, phenotypic data used in mapping and the selection process should encompass all three dimensions of the seed.

Two general considerations for plant breeding include amount of data generated per genotype and reduction of breeding cycle time. First, even with increases in the amount data extracted from each genotype, the overall number of evaluated lines has been stable or increasing. Reduction of breeding cycle time has been achieved by using off-season nurseries, doubled-haploids, and phenotypic prediction methodologies (Peleman and van der Voort 2003; Eathington *et al.* 2007; Heffner *et al.* 2010). The increase in the intensity of plant breeding efforts highlights the need for familiarity with high throughput phenotyping techniques in breeding programs of all sizes. Exploring fine-scale morphological data using DIA, as described here, is consistent with these trends and may allow for more directed manipulation of agronomic and quality traits of wheat in the future.

Exploratory studies such as this are limited by the amount of time required for phenotyping. The imaging process reported here took approximately seven minutes per line, not including post-processing steps prior to data analysis. If the assumption is made that most genetic studies will use population sizes based on 96 or 384 well plate formats dictated by genotyping methods, and that several environments are required, a study with a population size of 384 lines replicated once per location, over three locations for three seasons would require ~800 hours of labor to capture the kernel morphology images. Because of this time requirement, selection experiments will use machines that measure seed only in the horizontal dimension and report thickness values from prolate spheroid geometry, or other derived metrics, without actually measuring thickness. For items that are uniformly round such approximate measures may work well, but for those that are less uniform in shape (such as wheat seeds, which have a crease), measures based on assumptions of radial symmetry are less accurate for measuring volume. As demonstrated here, the shape of wheat seeds when photographed with each kernel on end is able to detect a significant QTL on chromosome 7D co-localized with a flour

yield QTL that was not revealed by measuring length or width alone. This demonstrates the utility of three-dimensional imaging for phenotyping.

In this study, the two DIA approaches to phenotyping kernel morphology were used for (1) characterizing correlations between different seed shape descriptors and flour yield as well as (2) detecting QTL and clarifying the role of the underlying genes in affecting overall shape of the seed. Additionally, the entire three dimensional shape of the seed described using two images in different orientations was shown to identify seed shape QTL that co-locate with flour yield and would go undetected based a single two dimensional image of the seed. We reported QTL for length, width, and vertical perimeter that were co-localized with QTL for flour yield. Finally, we provided guidelines for implementation of DIA using both measures of major dimensions and fine-scale EFD measures of shape. These results will facilitate the dissection of the genetic relationships influencing the shape of not only seeds, but also other plant organs of economic importance.

Table 2.1. Phenotypes examined for comparison of digital image analysis methods.

Direct Photometric Measures, ImageJ	Abbreviation	Measured as:
Seed Length	LENGTH	Major axis of horizontal image
Seed Width	WIDTH	Minor axis of horizontal image (equivalent to major axis of vertical image)
Seed Thickness	THICK	Minor axis of vertical image
Horizontal Perimeter	HPERIM	Perimeter from horizontal image
Vertical Perimeter	VPERIM	Perimeter from vertical image
Horizontal Area	HAREA	Area from horizontal image
Vertical Area	VAREA	Area from vertical image
Direct Photometric Measures, SHAPE	**Abbreviation**	**Measured as:**
Horizontal PC1	HPC1	See Figure 2.4
Horizontal PC2	HPC2	See Figure 2.4
Horizontal PC3	HPC3	See Figure 2.4
Horizontal PC4	HPC4	See Figure 2.4
Horizontal PC5	HPC5	See Figure 2.4
Vertical PC1	VPC1	See Figure 2.5
Vertical PC2	VPC2	See Figure 2.5
Vertical PC3	VPC3	See Figure 2.5
Vertical PC4	VPC4	See Figure 2.5
Vertical PC5	VPC5	See Figure 2.5
Derived Photometric Measures	**Abbreviation**	**Derivation:**
Aspect Ratio	ASPECT	(LENGTH) / (WIDTH)
Factor Form Density	FFD	individual grain weight / (LENGTH * WIDTH)
Volume	VOL_{XYZ}	(4/3)π(x)(y)(z), where x, y, & z represent LENGTH, WIDTH, & THICK
Deviation from an optimal ellipse, horizontal	PDEVH	PDEVH = |(p – HPERIM) / HPERIM | where p = π [3(LENGTH + WIDTH) – √ (3*LENGTH + WIDTH)*(LENGTH + 3*WIDTH)
Deviation from an optimal ellipse, vertical	PDEVV	PDEVV = |(p – VPERIM) / VPERIM | where p = π [3(THICK + WIDTH) – √ (3*THICK + WIDTH)*(THICK + 3*WIDTH)
Composite 1	COMP1	(VOL_{XYZ} / PDEVH) / (PDEVV)
Composite 2a	COMP2a	(VOL_{XYZ})*(HPC1)*(VPC2)
Composite 2b	COMP2b	(VOL_{XYZ})*(HPC1)*(VPC3)
Composite 2c	COMP2c	(VOL_{XYZ})*(HPC1)*(VPC4)
Other Phenotypes	**Abbreviation**	**Measured as:**
Thousand Kernel Weight	TKW	weight of individual kernel from variable number seed sample multiplied by 1,000
Flour Yield	FLYLD	standardized measure of flour yield from USDA soft wheat quality lab

Table 2.2, following page. Pearson's Correlation Coefficient values including direct phenotypic measures of seed morphology as well as flour yield in Cayuga X Caledonia. Abbreviations for individual environments are as follows: _sny05 = Snyder 2005, _sny06 = Snyder 2006, _sny08 = Snyder 2008 _helf = Helfer 2005, _mcg = McGowan 2005, _ket = Ketola 2005. Correlations were tested for significance by comparing observed r-values to critical r-values of 0.13 and 0.15 for a two-tailed test at significance levels of 0.05 (*) and 0.01 (**) for df = 159.

Table 2.2 (continued)

= 0.10 – 0.25
= 0.26 – 0.50
= 0.51 – 0.75
= 0.76 – 1.00

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. FLYLD_sny05															
2. FLYLD_sny06	0.51**														
3. FLYLD_sny08	0.48**	0.43**													
4. LENGTH_helf	0.09	0.12	-0.02												
5. LENGTH_mcg	0.05	0.16**	-0.04	0.74**											
6. LENGTH_ket	0.04	0.12	-0.04	0.83**	0.72**										
7. WIDTH_helf	-0.18**	0.19**	0.15**	0.11	0.12	0.08**									
8. WIDTH_mcg	0.03	0.03	0.06	0.07	0.37**	0.11	0.49**								
9. WIDTH_ket	-0.13*	0.11	0.07	0.02	0.08	0.12	0.49**	0.49**							
10. THICK_helf	0.04	0.02	0.06	0.12	0.03	0.03	0.54**	0.43**	0.33**						
11. THICK_mcg	0.08	0.14*	-0.07	0.03	0.35**	0.11	0.30**	0.64**	0.36**	0.50**					
12. THICK_ket	0.00	0.07	-0.01	0.01	0.05	0.08	0.20**	0.32**	0.69**	0.36**	0.51**				
13. HPC1_helf	-0.16**	0.00	0.17**	-0.67**	-0.51**	-0.59**	0.47**	0.29**	0.37**	0.27**	0.14*	0.20**			
14. HPC1_mcg	-0.13*	-0.09	0.02	-0.71**	-0.62**	-0.65**	0.33**	0.36**	0.38**	0.28**	0.25**	0.25**	0.81**		
15. HPC1_ket	-0.16*	-0.03	0.07	-0.68**	-0.55**	-0.69**	0.31**	0.26**	0.52**	0.23**	0.14*	0.30**	0.79**	0.84**	
16. HPC2_helf	0.12	0.06	0.17**	-0.09	-0.22**	-0.07	0.07	0.00	0.09	0.08	-0.03	0.09	0.13*	0.13*	0.09
17. HPC2_mcg	0.14*	0.00	0.24**	-0.10	-0.35**	-0.11	-0.08	-0.14*	-0.11	-0.06	-0.38**	-0.10	-0.03	-0.04	-0.01
18. HPC2_ket	0.12	0.09	0.18**	-0.17**	-0.24**	-0.09	0.31**	-0.09	0.08	-0.05	-0.01	0.19**	0.08	0.09	0.07
19. HPC3_helf	-0.02	0.06	0.02	0.50**	0.34**	0.34**	0.01	-0.03	-0.01	-0.03	-0.09	-0.07	-0.27**	-0.32**	-0.26**
20. HPC3_mcg	-0.02	-0.06	0.05	0.38**	0.17**	0.26**	-0.05	0.01	-0.08	-0.01	-0.33**	-0.18**	-0.26**	-0.28**	-0.23**
21. HPC3_ket	0.04	0.13*	0.02	0.48**	0.49**	0.54**	0.02	0.06	0.10	0.03	0.10	0.10	-0.32**	-0.35**	-0.27**
22. VPC2_helf	0.21**	-0.04	-0.01	0.18**	0.11	0.11	-0.17**	0.06	-0.14*	0.41**	0.15**	0.07	-0.26**	-0.12	-0.19**
23. VPC2_mcg	0.01	0.19**	-0.08	0.03	0.32**	0.12	0.13*	0.12	0.14*	0.18**	0.68**	0.29**	0.03	0.06	0.02
24. VPC2_ket	0.01	-0.03	-0.16**	-0.07	0.01	0.04	-0.09	0.03	-0.26**	0.01	0.24**	-0.01	-0.07	0.00	-0.17**
25. VPC3_helf	-0.09	0.18**	0.15**	0.13*	0.17**	0.14*	0.32**	-0.01	0.13*	-0.30**	-0.17**	-0.12	0.04	-0.07	-0.03
26. VPC3_mcg	0.03	0.17**	0.03	0.09	0.39**	0.17**	0.16**	0.20**	0.20**	-0.13*	0.25**	0.06	0.02	0.00	0.02
27. VPC3_ket	-0.06	0.14*	0.09	0.13*	0.20**	0.08	0.10	0.09	0.21*	-0.21**	-0.16**	-0.26**	-0.08	-0.10	0.03
28. VPC4_helf	-0.14*	0.24**	0.08	-0.02	0.16**	0.10	0.15**	0.00	0.10	-0.28**	0.00	-0.06	0.09	0.01	0.07
29. VPC4_mcg	-0.10	0.16**	-0.06	-0.08	0.26**	0.04	0.19**	0.16**	0.24**	0.00	0.44**	0.23**	0.17**	0.13*	0.18**
30. VPC4_ket	-0.08	0.03	-0.07	-0.12	-0.07	-0.02	0.09	-0.02	-0.14*	-0.15**	-0.16**	-0.47**	0.09	0.04	0.04

	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
16. HPC2_helf															
17. HPC2_mcg	0.57**														
18. HPC2_ket	0.57**	0.58**													
19. HPC3_helf	-0.23**	-0.11	-0.09												
20. HPC3_mcg	0.01	0.27**	-0.08	0.47**											
21. HPC3_ket	-0.25**	-0.24**	-0.18**	0.57**	0.40**										
22. VPC2_helf	-0.19**	-0.08	-0.17**	0.02	0.16**	0.16**									
23. VPC2_mcg	-0.18**	-0.61**	-0.06	0.01	-0.56**	0.20**	0.04								
24. VPC2_ket	0.05	-0.13*	-0.12	-0.18**	-0.08	-0.20**	0.07	0.18**							
25. VPC3_helf	-0.16**	-0.19**	-0.14*	-0.02	-0.09	0.08	-0.36**	0.09	-0.06						
26. VPC3_mcg	-0.23**	-0.51**	-0.13*	0.10	-0.50	0.21**	-0.15*	0.68**	-0.06	0.42**					
27. VPC3_ket	-0.10	-0.15**	-0.26**	0.09	0.15**	0.01	-0.13	-0.08	0.02	0.42**	0.23**				
28. VPC4_helf	0.09	-0.07	0.08	0.01	-0.03	0.23**	-0.42**	0.21**	-0.03	0.35**	0.33**	0.09			
29. VPC4_mcg	-0.12	-0.53**	-0.02	-0.01	-0.55**	0.13*	-0.19**	0.78**	0.00	0.24**	0.77**	-0.02	0.50**		
30. VPC4_ket	0.09	0.06	-0.09	-0.06	0.09	-0.14*	-0.27**	-0.13*	0.24**	0.08	-0.05	0.16**	0.35**	0.09	

Table 2.3, following page. Pearson's Correlation Coefficient values including derived phenotypic measures of seed morphology as well as flour yield in Cayuga X Caledonia. Abbreviations for individual environments are as follows: _sny05 = Snyder 2005, _sny06 = Snyder 2006, _sny08 = Snyder 2008 _helf = Helfer 2005, _mcg = McGowan 2005, _ket = Ketola 2005. . Correlations were tested for significance by comparing observed r-values to critical r-values of 0.13 and 0.15 for a two-tailed test at significance levels of 0.05 (*) and 0.01 (**) for df = 159.

Table 2.3 (continued)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. FLYLD_sny05															
2. FLYLD_sny06	0.51**														
3. FLYLD_sny08	0.48**	0.43**													
4. ASPECT_helf	0.19**	-0.01	-0.11												
5. ASPECT_mcg	0.03	0.15**	-0.08	0.74**											
6. APSECT_ket	0.13*	0.02	-0.07	0.73**	0.72**										
7. FFD_helf	0.04	0.06	0.07	-0.34**	-0.31**	-0.35**									
8. FFD_mcg	-0.13*	0.04	-0.01	-0.35**	-0.33**	-0.37**	0.50**								
9. FFD_ket	-0.09	0.20**	0.06	-0.30**	-0.19**	-0.26**	0.37**	0.42**							
10. VOL_{XYZ}_helf	0.00	0.16**	0.09	0.07	0.20**	0.20**	0.12	0.20**	0.28**						
11. VOL_{XYZ}_mcG	0.09	0.15**	-0.01	0.12	0.19**	0.14*	0.11	0.10	0.25**	0.55**					
12. VOL_{XYZ}_ket	-0.01	0.14*	0.00	0.14*	0.12	-0.11	0.16**	0.19**	0.04	0.43**	0.49**				
13. PDEVH_helf	0.24**	0.08	0.00	-0.07	0.02	0.05	0.06	-0.04	0.04	-0.12	0.00	-0.24**			
14. PDEVH_mcg	0.19**	0.08	-0.05	-0.02	0.17**	-0.04	-0.11	-0.17**	-0.14*	-0.13	0.05	0.07	0.25**		
15. PDEVH_ket	0.13*	0.18**	0.10	-0.13*	-0.04	-0.07	-0.11	-0.10	0.09	-0.08	-0.12	-0.28**	0.43**	0.21**	
16. PDEVV_helf	-0.03	-0.10	-0.17**	-0.07	0.01	0.03	-0.01	-0.11	-0.08	-0.14	-0.07	-0.21**	0.19**	0.12	0.08
17. PDEVV_mcg	-0.07	-0.10	-0.06	-0.05	0.15**	-0.10	0.02	0.06	-0.08	-0.09	-0.16**	0.08	-0.05	0.34**	-0.11
18. PDEVV_ket	0.03	-0.20**	-0.05	0.08	0.07	0.16**	-0.11	-0.02	0.03	0.08	-0.02	-0.17	0.11	-0.06	0.21**
19. COMPS1_helf	-0.11	0.06	0.11	0.07	0.03	0.01	0.02	0.15**	0.12	0.51**	0.22**	0.39**	-0.71**	-0.25**	-0.31**
20. COMPS1_mcg	-0.03	0.09	0.05	0.05	-0.10	0.11	0.04	0.07	0.18**	0.28**	0.40**	0.06	-0.08	-0.63**	-0.05
21. COMPS1_ket	-0.04	0.03	0.00	0.07	-0.01	-0.09	0.15**	0.11	-0.07	0.12	0.21**	0.56**	-0.41**	-0.07	-0.81**
22. COMPS2_helf	-0.07	-0.05	-0.01	-0.54**	-0.47**	-0.40**	0.20**	0.15	0.15**	-0.20**	-0.10	-0.07	-0.06	0.03	-0.10
23. COMPS2_mcg	-0.12	0.17**	0.08	-0.35**	0.56**	0.34**	-0.18**	-0.20**	-0.05	0.18**	0.18**	0.07	0.06	0.22**	0.00
24. COMPS2_ket	-0.12	0.05	0.07	-0.37**	-0.37**	-0.51**	0.23**	0.20**	0.19**	-0.06	-0.07	-0.03	-0.03	-0.10	-0.06
25. COMPS2b_helf	-0.13*	-0.05	-0.03	0.16**	0.18**	0.14*	-0.19**	-0.11	-0.13	0.01	-0.04	-0.12	0.04	-0.05	0.07
26. COMPS2b_mcg	0.09	0.17**	0.21**	0.12	0.33**	0.15*	-0.12	-0.11	-0.01	0.11	0.17**	0.03	0.05	0.23**	0.02
27. COMPS2b_ket	-0.05	-0.11	0.12	-0.25**	-0.19**	-0.17**	0.05	0.03	0.05	-0.09	0.00	-0.18**	0.03	-0.08	0.05
28. COMPS2c_helf	-0.04	0.02	0.19**	-0.64**	-0.51**	-0.54**	0.23**	0.30**	0.15**	-0.04	-0.11	-0.07	-0.09	0.02	0.10
29. COMPS2c_mcg	0.14*	0.19**	0.15**	0.14*	0.38**	0.15*	-0.10	-0.12	0.04	0.14*	0.17**	0.03	0.05	0.27**	0.04
30. COMPS2c_ket	-0.07	0.08	0.02	-0.01	-0.03	-0.05	-0.10	-0.01	0.03	0.08	0.04	0.01	-0.06	-0.07	0.12

	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
16. PDEVV_helf															
17. PDEVV_mcg	0.05														
18. PDEVV_ket	0.28**	0.12													
19. COMPS1_helf	-0.71**	-0.05	-0.19**												
20. COMPS1_mcg	-0.13	-0.84**	-0.04	0.22**											
21. COMPS1_ket	-0.24**	0.05	-0.61**	0.41**	0.05										
22. COMPS2_helf	-0.04	0.02	-0.16**	0.01	-0.05	0.14*									
23. COMPS2_mcg	0.01	0.12	0.03	0.01	-0.12	0.01	-0.25**								
24. COMPS2_ket	-0.02	0.05	-0.02	0.01	-0.01	0.04	0.23**	-0.12							
25. COMPS2b_helf	0.08	0.07	0.01	-0.06	-0.03	-0.09	-0.38**	0.18**	-0.09						
26. COMPS2b_mcg	-0.01	0.21**	-0.03	0.02	-0.19**	0.03	-0.17**	0.71**	-0.17**	0.30**					
27. COMPS2b_ket	0.13*	-0.07	0.04	-0.13**	0.07	-0.12	0.13*	-0.26**	0.05	0.30**	0.05				
28. COMPS2c_helf	-0.05	0.11	-0.05	0.08	-0.09	-0.06	0.16*	-0.16**	0.34**	0.11	0.03	0.22**			
29. COMPS2c_mcg	0.01	0.19**	0.04	0.02	-0.18**	-0.02	-0.24**	0.84**	-0.04	0.31**	0.83**	-0.10	0.12		
30. COMPS2c_ket	-0.03	0.03	0.12	0.08	0.03	-0.17**	-0.22**	-0.10	0.28**	0.05	-0.11	0.13*	0.34**	0.06	

= 0.10 – 0.25
= 0.26 – 0.50
= 0.51 – 0.75
= 0.76 – 1.00

Table 2.4. Summary of kernel morphology QTL.

Trait	Chromosome	r^2	LOD	Confidence Interval (cM)	Bordering Markers	Nearest Marker
LENGTH	2A	0.08	4.8	9cM	*E40M59257Y, Xgwm526b*	*E40M59257Y*
	3A	0.08	6.3	25cM	*Xwmc264, Xbarc0045*	*Xbarc0346*
	3B	0.16	7.7	20cM	*Xbarc0077, E35M49161L*	*Xbarc0077*
	3D	0.21	18	24cM	*Xcfd70, Xgwm645*	*Xbarc1161*
	5D	0.10	5.5	37cM	*Xgwm639b, E36M60270L*	*Xwmc150a*
	6D	0.10	8	14cM	*Xbarc0204, Xbarc0096*	*Xbarc0204*
WIDTH	2A	0.09	3.7	3cM	*Xgwm339, Xwmc522*	*Xgwm339*
	2B	0.08	4.2	9cM	*Xbarc0328, Xwmc360*	*wPt-2430*
	4B	0.07	5	18cM	*Xwmc435, Xbarc0163*	*Xgwm192*
	4D	0.19	8	30cM	*Xbarc0217, Xbarc1118*	*RHT-DF-MR2*
	6D	0.07	4.4	14cM	*Xbarc0096, Xcfd01*	*Xbarc0196*
THICK	2A	0.16	5.9	20cM	*Xgwm339, Xgwm095*	*Xgwm339, Xgwm515*
	3B	0.08	5.4	25cM	*E35M49161L, TRAP_telos12-32*	*Xbarc0229*
	4D	0.10	5	24cM	*Xbarc0217, Xbarc1118*	*RHT-DF-MR2*
HAREA	2B	0.08	3.9	5cM	*Xwmc360, wPt9190*	*Xwmc360*
	3D	0.13	9	22cM	*Xcfd64, Xbarc1161*	*Xcfd70*
	6D	0.17	10	36cM	*Xbarc0096, Xcfd01*	*Xcfd37*
VAREA	2A	0.08	4.9	10cM	*wPt-6361, Xwmc522*	*Xgwm339*
	2A	0.08	4.2	2.5cM	*Xwmc522, Xgwm515*	*Xgwm515*
	4D	0.10	4.6	19cM	*Xbarc0217, Xbarc1118*	*RHT-DF-MR2*
HPERIM	3D	0.08	7.2	29cM	*Xcfd64, gpw4152*	*Xcfd70*
	5D	0.07	4	15cM	*Xgwm639b, Xwmc150a*	*Xwmc150a*
	6D	0.12	6.2	28cM	*Xbarc0096, Xcfd01*	*Xcfd37*
VPERIM	2A	0.07	3.7	3cM	*Xgwm339, Xwmc522*	*Xgwm339*
	2B	0.08	4.8	19cM	*Xgwm429, wPt2430*	*Xwmc474, Xbarc0328*
	7D	0.15	6.1	27cM	*Xwmc671, Xbarc0172*	*Xwmc150b*
HPC1	2B	0.07	7	20cM	*Xwmc770, Xbarc0328*	*Xgwm429*
	3A	0.18	16	25cM	*Xwmc264, Xbarc0045*	*Xbarc0346*
	3B	0.09	11	20cM	*Xbarc0077, E35M49161L*	*Xbarc0077*
	3D	0.18	10.8	24cM	*Xcfd70, Xgwm645*	*Xbarc1161*
	4D	0.09	11	35cM	*Xbarc0217, Xbarc1118*	*RHT-DF-MR2*
	5D	0.07	6	49cM	*Xgwm639b, E36M60270L*	*Xwmc150a*
	6D	0.07	5.5	8cM	*Xbarc0204, Xbarc0096*	*Xbarc0204*
HPC2	2D	0.14	6.8	29cM	*wPt-9997, wPt-4144*	*Xbarc1123*
HPC3	3B	0.16	6.6	31cM	*Xbarc0077, Xbarc0229*	*E34M49161L*
	3D	0.08	7.2	17cM	*Xcfd70, gpw4125*	*Xbarc1161*
VPC2	-	-	-	-	-	-
VPC3	1A	0.07	4.5	10cM	*Xbarc0028, E42M49146L*	*CFA2129_RTL*
	2B	0.08	4.3	16cM	*Xbarc0328, wPt-9190*	*wPt-2430, Xwmc360*
VPC4	6D	0.2	9.3	42cM	*Xbarc0096, Xcfd01*	*Xcfd37*
TKW	4D	0.14	6.5	29cM	*Xbarc0217, Xbarc1118*	*RHT-DF-MR2*
	2B	0.10	5.2	16cM	*Xbarc0328, wPt-9190*	*Xwmc360*
	7D	0.09	4.6	10cM	*Xwmc150b, Xbarc172*	*Xwmc150b*
ASPECT	3A	0.05	10	25cM	*Xwmc264, Xbarc0045*	*Xbarc0346*
	3B	0.13	7	20cM	*Xbarc0077, E35M49161L*	*Xbarc0077*
	3D	0.18	11.5	24cM	*Xcfd70, Xgwm645*	*Xbarc1161*
	4D	0.12	7.2	26cM	*Xbarc217, Xbarc1118*	*RHT-DF-MR2*
	6D	0.05	4	5cM	*Xbarc0204, Xbarc0096*	*Xbarc0204*
FFD	4D	0.2	10	27cM	*Xbarc0217, Xbarc1118*	*RHT-DF-MR2*
VOL_{XYZ}	3D	0.08	3.1	2cM	*Xbarc0006, Xwmc294*	*Xcfd70*
PDEVH	2B	0.07	8	28cM	*Xwmc770, wPt02430*	*Xwmc474*
	3D	0.08	5	20cM	*Xwmc294, Xgwm645*	*gpw4152*
PDEVV	-	-	-	-	-	-
COMP1	3D	0.09	4.8	16cM	*Xbarc0125, Xgwm645*	*gpw4152*
	4D	0.16	7	32cM	*Xbarc0217, Xbarc1118*	*RHT-DF-MR2*
COMP2a	-	-	-	-	-	-
COMP2b	-	-	-	-	-	-
COMP2c	-	-	-	-	-	-
FLYLD	2B	0.08	4.2	10.9cM	*Xbarc0328, wPt9190*	*wPt2430*
	3A	0.09	3.6	8.6cM	*Xwmc264, Xbarc0045*	*Xbarc0346*
	7D	0.08	3.9	5.9cM	*Xwmc150b, Xgwm473*	*Xbarc0172*

Table 2.5. Summary of average r^2, LOD, confidence interval (CI) values and distribution of QTL based on phenotyping method.

					Number of QTL detected per chromosome																				
	Total QTL	Avg r^2	Avg LOD	Avg CI	1A	1B	1D	2A	2B	2D	3A	3B	3D	4A	4B	4D	5A	5B	5D	6A	6B	6D	7A	7B	7D
ImageJ	**26**	**0.11**	**6.17**	**18.8**				6	3		1	2	3		1	3			2			4			1
SHAPE	**13**	**0.10**	**8.15**	**25**	1				2	1	1	2	2			1			1			2			
Derived (including TKW)	**14**	**0.11**	**6.71**	**20.4**			1		1			1	4			4						2			1
TKW	**3**	**0.11**	**5.4**	**20**					1							1									1

Table 2.6. Co-localization of QTL detected by ImageJ and SHAPE.

Chromosome	QTL detected for	includes	does not include
2A	VPERIM, WIDTH, VAREA, THICK	width, thickness	length, flour yield
2A	HPERIM, LENGTH	length	width, thickness, flour yield
2B	HAREA, WIDTH, VPC3, VPERIM, HPC1, FLYLD	width, flour yield	length, thickness
3A	LENGTH, HPC1, FLYLD	length, flour yield	width, thickness
3B	HPC3, THICK, LENGTH, HPC1	length, thickness	width, flour yield
3D	HAREA, HPERIM, LENGTH, HPC3, HPC1	length	width, thickness, flour yield
4D	VAREA, THICK, WIDTH, HPC1	thickness, width	length, flour yield
6D	WIDTH, HPC1, HPERIM, LENGTH, VPC4, HAREA	length, width	thickness, flour yield
7D	VPERIM, FLYLD	VPERIM, flour yield	length, width

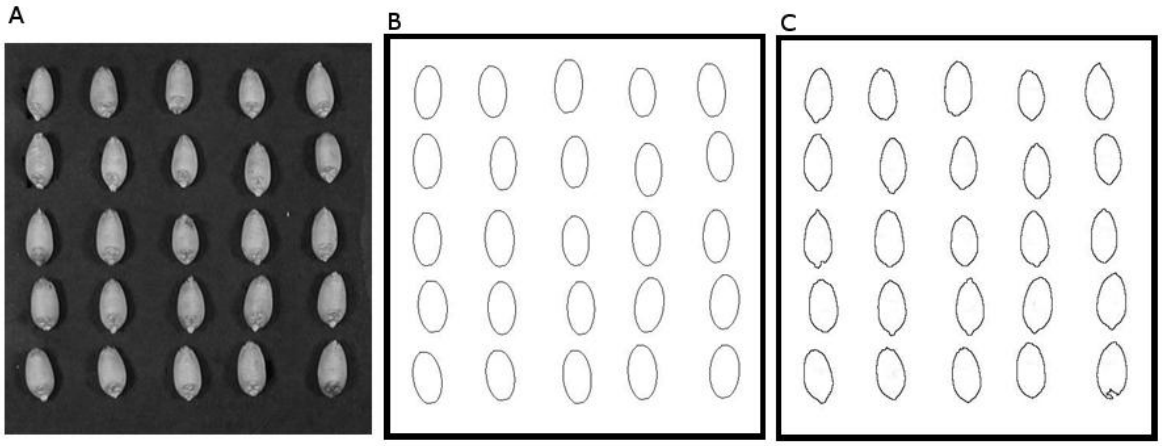

Figure 2.1. Horizontal image (H image) processing via ImageJ with original photograph pictured in A, subsequent transformation to fitted ellipses in B, and outlines used in checking for poor quality measurements shown in C.

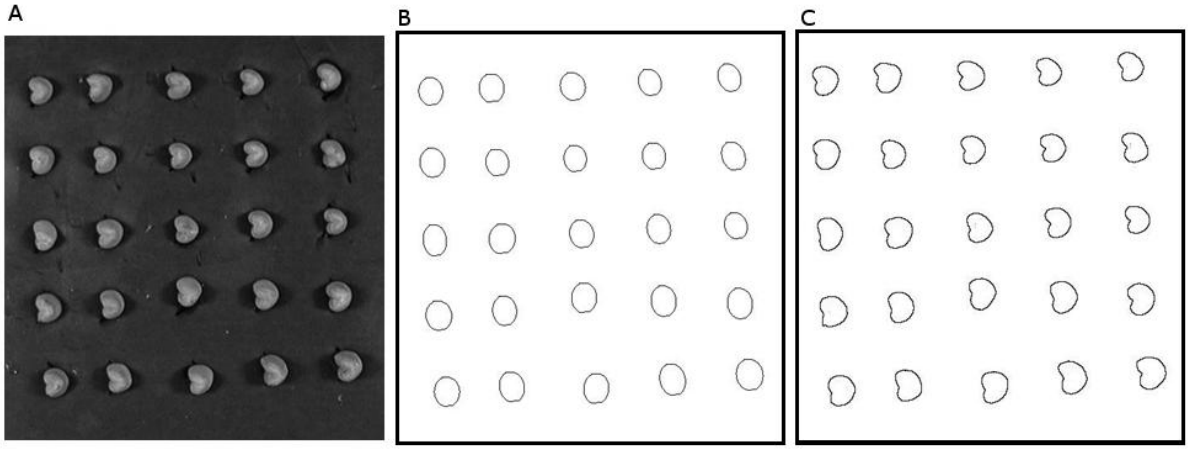

Figure 2.2. Vertical image (V image) processing via ImageJ with original photograph pictured in A, subsequent transformation to fitted ellipses in B, and outlines used in checking for poor quality measurements shown in C.

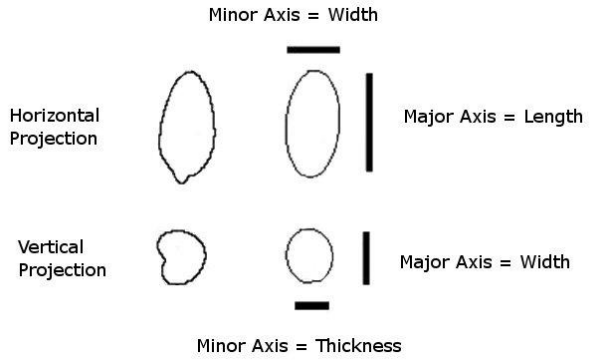

Figure 2.3. Conversion of seed images into axes measurements via ImageJ.

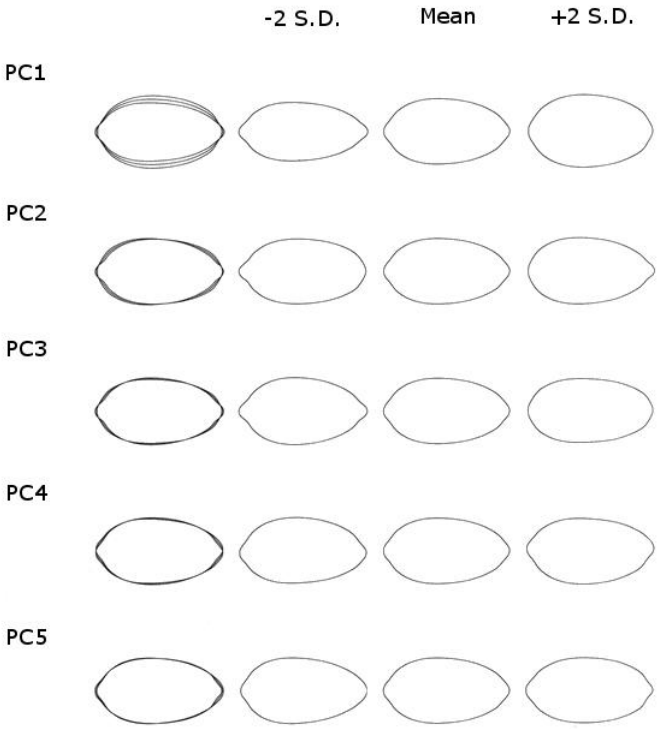

Figure 2.4. Principle components returned from SHAPE for horizontal images of Cayuga X Caledonia seeds.

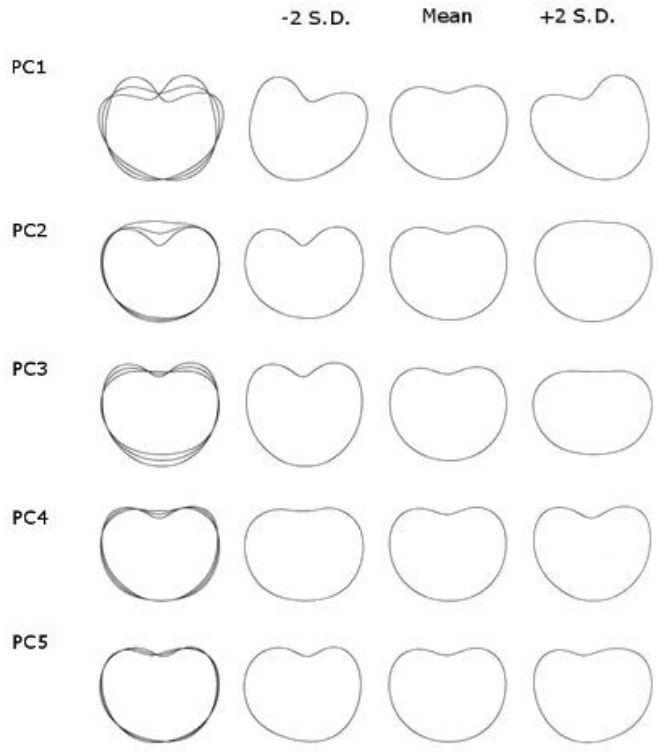

Figure 2.5. Principle components returned from SHAPE for vertical images of Cayuga x Caledonia seeds.

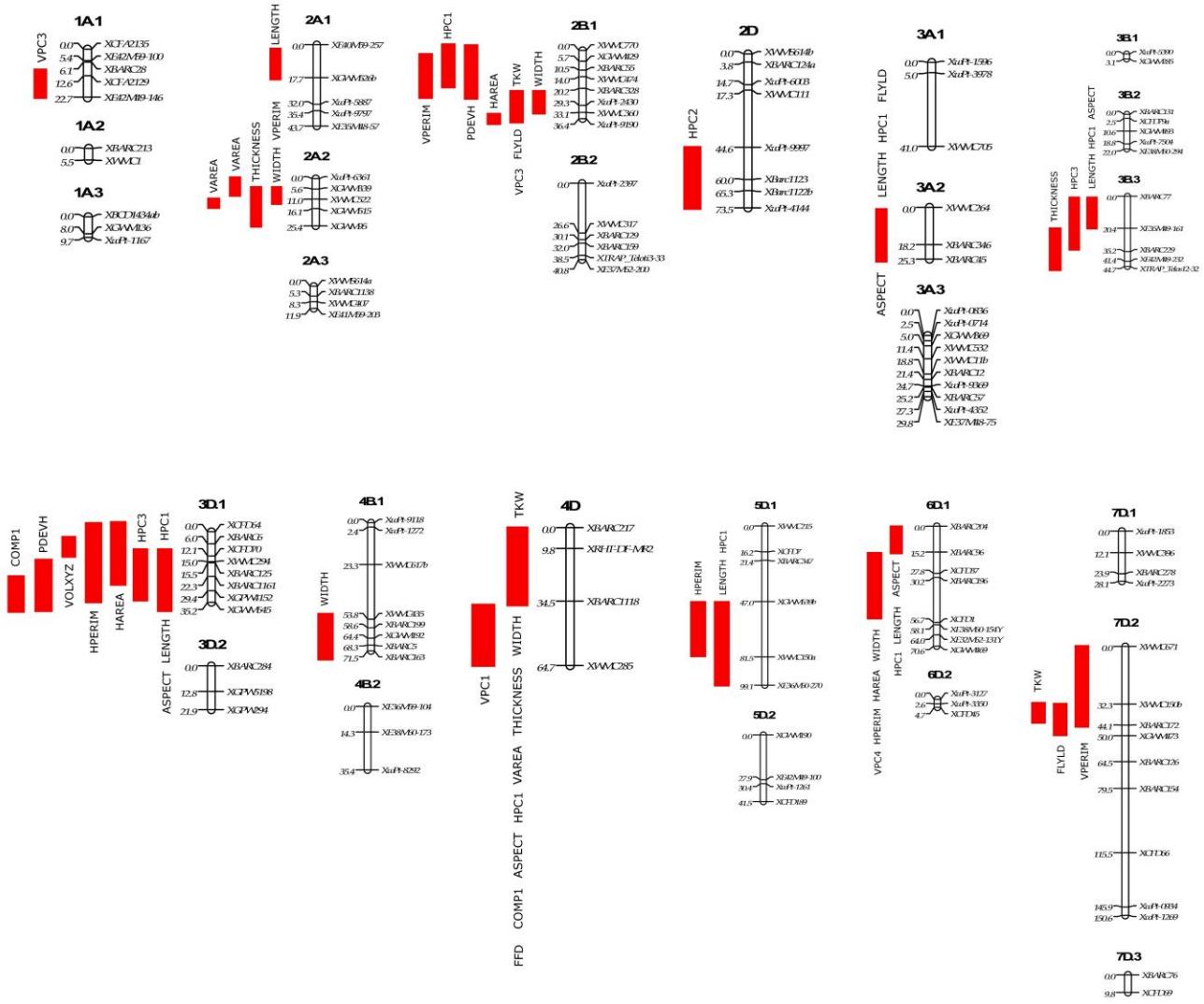

Figure 2.6. Graphical representation of significant QTL detected in Cayuga x Caledonia, with QTL regions denoted by red bars and traits affected by each QTL listed to either side.

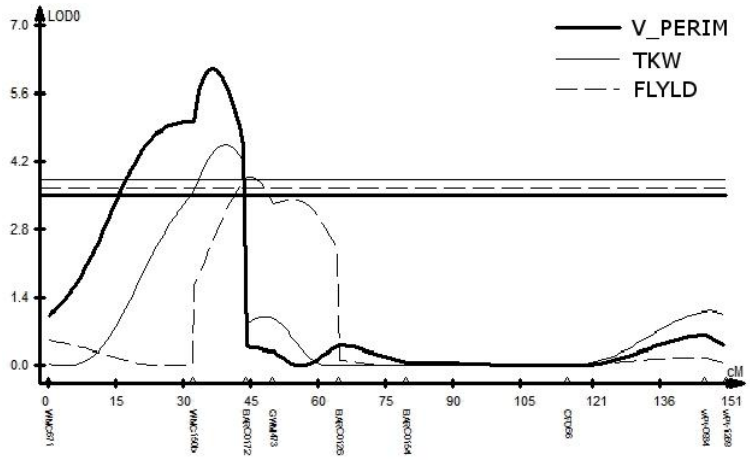

Figure 2.7. Co-localization of QTL for FLYLD, TKW, and V_PERIM on chromosome 7D.

REFERENCES

Andrews L (2002) Quality Characteristics of Soft Wheat Cultivars. USDA ARS Soft Wheat Quality Laboratory, Wooster, OH

Barnard AD, Labuschagne MT, van Nierkerk HA (2002) Heritability estimates of bread wheat quality traits in the Western Cape province of South Africa. Euphytica 127:115-122

Breseghello F, Sorrels ME (2005) Genetic loci related to kernel quality differences between a soft and a hard wheat cultivar. Crop Sci 45:1685-1695

Breseghello F, Sorrells ME (2006) Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. Genetics 172:1165-1177

Breseghello F, Sorrells ME (2007) QTL analysis of kernel size and shape in two hexaploid wheat mapping populations. Field Crops Research 101:172-179

Campbell KG, Bergman CJ, Gualberto DG, Anderson JA, Giroux MJ, Hareland G, Fulcher RG, Sorrells ME, Finney PL (1999) Quantitative trait loci associated with kernel traits in a soft x hard wheat cross. Crop Sci 39:1184-1195

Cober ER, Voldeng HD, Fregeau-Reid JA (1997) Heritability of seed shape and seed size in soybean. Crop Sci 37:1767-1769

Dana W, Ivo W (2008) Computer image analysis of seed shape and seed color for flax cultivar description. Comput Electron Agr 61:126-135

DeSouza N (2010) High-throughput phenotyping. Nat Methods 7:36

Dholakia BB, Ammiraju JSS, Singh H, Lagu MD, Roder MS, Rao VS, Dhaliwal HS, Ranjekar PK, Gupta VS (2003) Molecular marker analysis of kernel size and shape in bread wheat. Plant Breeding 122:392-395

Diao X, Furuno T, Fujita M (1999) Digital image analysis of cross-sectional tracheid shapes in Japanese softwoods using the circularity index and aspect ratio. J Wood Sci 45:98-105

Doehlert DC, McMullen MS, Jannink J, Panigrahi S, Gu H, Riveland NR (2004) Evaluation of oat kernel size uniformity. Crop Sci 44:1178-1186

Eathington SR, Crosbie TM, Edwards MD, Reiter RS, Bull JK (2007) Molecular markers in a commercial breeding program. Crop Sci 47:S-154

Finney PL, Andrews LC (1986) Revised microtesting for soft wheat quality evaluation. Cereal Chem 63:177-182

Gegas V, Nazari A, Griffiths S, Simmonds J, Fish L, Orford S, Sayers L, Doonan JH, Snape JW (2010) A genetic framework for grain size and shape and shape variation in wheat. The Plant Cell 22:1046-1056

Giura A, Saulescu NN (1996) Chromosomal location of genes controlling grain size in a large grained selection of wheat (*Triticum aestivum* L.). Euphytica 89:77-80

Goto S, Iwata H, Shibano S, Ohya K, Suzuki A, Ogawa H (2005) Fruit shape variation in *Fraxinus mandshurica* var. *japonica* characterized using elliptic Fourier descriptors and the effect on flight duration. Ecol Res 20:733-738

Guo L, Ma L, Jiang H, Zeng D, Hu J, Wu L, Gao Z, Zhang G, Qian Q (2009) Genetic analysis and fine mapping of two genes for grain shape and weight in rice. J Integr Plant Biol 51:45-51

Heffner EL, Lorenz AJ, Jannink J, Sorrells ME (2010) Plant breeding with genomic selection: gain per unit time and cost. Crop Sci 50:1681-1690

Himstedt M, Fricke T, Wachendorf M (2009) Determining the contribution of legumes in legume-grass mixtures using digital image analysis. Crop Sci 49:1910-1916

Ibaraki Y, Kenji K (2001) Application of image analysis to plant cell suspension cultures. Comput Electron Agr 30:193-203

Horgan GW (2001) The statistical analysis of plant part appearance — a review. Comput Electron Agr 31:169-190

Houle D, Govindaraju DR, Omholt S (2010) Phenomics: the next challenge. Nat Rev Genet 11:855-866
Ibaraki Y, Kenji K (2001) Application of image analysis to plant cell suspension cultures. Comput Electron Agr 30:193-203

Iwata H, Niikura S, Matsuura S, Takano Y, Ukai Y (1998) Evaluation of variation of root shape of Japanese radish (*Raphanus sativus* L.) based on image analysis using elliptic Fourier descriptors. Euphytica 102:143-149

Iwata H, Ukai Y (2002) SHAPE: A computer program package for quantitative evaluation of biological shapes based on elliptic Fourier descriptors. J Hered 93:384-385

Iwata H, Ebana K, Uga Y, Hayashi T, Jannink J (2010) Genome-wide association study of grain shape variation among *Oryza sativa* L. germplasm based on elliptic Fourier analysis. Mol Breeding 25:203-215

Kozak M, Madry W (2006) Note on yield component analysis. Cereal Res Commun 34:933-940

Kwack MS, Kim EN, Lee H, Kim J, Chun S, Kim KD (2005) Digital image analysis to measure lesion area of cucumber anthracnose by *Colletotrichum orbiculare*. J Gen Plant Pathol 71:418-421

Marshall D, Mares D, Moss H, Ellison F (1986) Effects of grain shape and size on milling yields in wheat. II. Experimental studies. Aust J Agr Res 37:331-342

Marshall D, Ellison F, Mares D (1984) Effects of grain shape and size on milling yields in wheat. I. Theoretical analysis based on simple geometric models. Aust J Agr Res 35:619-630

Mohler V, Lukman R, Ortiz-Islas S, William M, Worland AJ, Beem JV, Wenzel G (2004) Genetic and physical mapping of photoperiod insensitivity gene *Ppd-B1* in common wheat. Euphytica 138:33-40

Montes JM, Melchinger AE, Reif JC (2007) Novel throughput phenotyping platforms in plant genetic studies. Trends Plant Sci 12:433-436

Munkvold JD, Tanaka JD, Benscher D, Sorrells ME (2009) Mapping quantitative trait loci for preharvest sprouting resistance in white wheat. Theor Appl Genet 119:1223-1235

Novaro P, Colucci F, Venora G, D'Egidio MG (2001) Image analysis of whole grains: A noninvasive method to predict semolina yield in durum wheat. Cereal Chem 78:217-221

Ohsawa R, Tsutsumi T, Uehara H, Namai H, Ninomiya S (1998) Quantitative evaluation of common buckwheat (*Fagopyrum esculentum* Moench) kernel shape by elliptic Fourier descriptor. Euphytica 101:175-183

Peleman JD, Van der Voort JR (2003) Breeding by design. Trends Plant Sci 8:330-334

Qiu X, Gong R, Tan Y, Yu S (2012) Mapping and characterization of the major quantitative trait locus *qSS7* associated with increased length and decreased width of rice seeds. Theor Appl Genet 125:1717-1726

Quarrie SA, Steed A, Calestani C, Semikhodskii A, Lebretron C, Chinoy C, Steele N, Plijevljakusi D, Waterman E, Weyen J, Schondelmaier J, Habash DZ, Farmer P, Sakerr L, Clarkson DT, Abugalieva A, Yessimbekova M, Turuspekov Y, Abugalieva S, Tuberosa R, Sanguineti MC, Hollington PA, Aragus R, Royo A, Dodiq D (2005) A high-density genetic map of hexaploid wheat (*Triticum aestivum* L.) from the cross Chinese Spring x Q1 and its use to compare QTLs for grain yield across a range of environments. Theor Appl Genet 110:865-880

Sadras VO, Egli DB (2008) Seed size variation in grain crops: allometric relationships between rate and duration of seed growth. Crop Sci 48:408-416

Shimoji H, Tokuda G, Tanaka Y, Moshiri B, Yamasaki H (2006) A simple method for two-dimensional color analyses of plant leaves. Russ J Plant Physl+ 53:126-133

Shouche SP, Rastogi R, Bhagwat SG, Sainis JK (2001) Shape analysis of grains of Indian wheat varieties. Comput Electron Agr 33:55-76

Shao G, Wei X, Chen M, Tang S, Luo J, Jiao G, Xie L, Hu P (2012) Allelic variation for a candidate gene for *GS7*, responsible for grain shape in rice. Theor Appl Genet 125:1303-1312

Somers D, Isaac P, Edwards K (2004) A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.). Theor Appl Genet 109:1105-1114

Sun XY, Wu K, Zhao Y, Kong FM, Han GZ, Jiang HM, Huang XJ, Li RJ, Wang HG, Li SS (2009) QTL analysis of kernel shape and weight using recombinant inbred lines in wheat. Euphytica 165:615-624

Tappan JH, Wright ME, Sistler FE (1987) Error sources in a digital image analysis system. Comput Electron Agr 2:109-118

Tsilo TJ, Hareland GA, Simsek S, Chao S, Anderson JA (2010) Genome mapping of kernel characteristics in hard red spring wheat breeding lines. Theor Appl Genet 121:717-730

Wang R, Hai L, Zhang X, You G, Yan C, Xiao S (2009) QTL mapping for grain filling rate and yield-related traits in RILs of the Chinese winter wheat population Heshangmai x Yu8679. Theor Appl Genet 118:313-325

White RJ, Prentice HC, Verwijst T (1988) Automated image acquisition and morphometric description. Can J Bot 66:450-459

Zhang LY, Liu DC, Guo XL, Yang WL, Sun JZ, Wang DW, Zhang A (2010) Genomic distribution of quantitative trait loci for yield and yield-related traits in common wheat. J Integr Plant Biol 62:996-100

CHAPTER THREE

Three dimensional seed size and shape QTL in hexaploid wheat (*Triticum aestivum* L.) populations

Abstract

Seed size and shape are important traits that have been altered during domestication and are monitored in wheat breeding programs because they may impact yield and milling quality. Two doubled-haploid (DH) wheat populations were used to map quantitative trait loci (QTL) for seed morphology using three seed axes and elliptic Fourier descriptors of shape. In the Synthetic W7984 x Opata M85 (SynOpDH) population, 50 QTL were detected on all chromosomes except 3D, 4D, 5D, 6B, and 7B. Seed shape QTL in SynOpDH were independent of genes *Q*, *Gpc-B1*, *Ser5B*, *S1*, *S2*, *S3*, and markers for rice *GW2* and *GS3*. Environmentally stable QTL on 1A and 2D, as well as a pleiotropic QTL on 5A were detected. QTL on chromosome 1A conditioned uniformly widened grains, 2D affected width, and 5A width and thickness. In the Cayuga x Caledonia DH population 32 QTL affecting kernel morphology were detected on chromosomes 1D, 2A, 2B, 3B, 3D, 4B, 4D, 6D, and 7D. The most significant QTL in this population were individual QTL on 3D affecting seed length (LOD 18, $r^2 = 0.21$), 1D affecting uniform widening (LOD 16, $r^2 = 0.18$), and 3B affecting uniform widening (LOD 10.8, $r^2 = 0.09$). Marker-assisted selection of these QTL for complex dimensions of seed shape may assist in breeding cultivars with improved yield or quality.

I. Introduction

Seed size and shape are important quantitative characteristics that are components of agronomic traits such as yield and milling quality of wheat cultivars. Although kernel morphology, notably seed size, has been researched since the 1950s the genes underlying the dimensions of wheat

seeds remain poorly understood and phenotypic selection of these traits to improve yield has met with little success (Yamazaki and Briggle 1969; Hook 1984; Parker *et al.* 1999). Earlier attempts at selecting larger (i.e., greater length & width) seeds improved flour extraction and grain protein concentration, but yield was unaffected (Wiersma *et al.* 2001). Selection for larger seed was accompanied by a reduction in the number of kernels per spike and number of tillers per plant. Due to such compensation of yield components, phenotypic selection for increases in seed size did not increase grain yield. Historical survey of breeding programs found little evidence to support a direct relationship between grain yield increases and changes in the number or size of seeds on an individual plant basis (Sayre *et al.* 1997). Yield increases were driven more by kernel number per square meter and harvest index. This suggests that complex physiological relationships may make it difficult to improve grain yield by manipulating component traits using only phenotypic data. Physiological trade-offs and allometric interactions between individual components of yield (kernel number, kernel weight, kernel shape, etc.) result in difficulty when using phenotypic data alone to select for specific components that could favorably impact yield.

Selection using individual dimensions of kernels has been proposed as a way to improve milling yield. Marshall *et al.* (1984) reviewed relevant literature on the relationship between kernel dimensions and milling yields. They proposed that better milling yields could be achieved by selecting for a larger, more spherical wheat kernel. Seeds with spherical shape and larger volume would have increased endosperm content relative to surface area. The increased endosperm content would presumably increase flour yield. However, their empirical results (Marshall *et al.* 1986) did not support theoretical predictions. The discrepancy between the theoretical and empirical studies was postulated to be due to a number of other unknown genetic factors that could affect milling quality (Marshall *et al.* 1986). A lack of relationship between kernel morphology and milling qualities has also been reported by other

research groups (Schuler *et al.* 1995; Bergman *et al.* 2000). The success of improving milling quality using component traits may require integration of genotypic data into the breeding process.

Numerous studies have reported on the molecular basis of phenotypic variation in wheat and how genotypic data can impact breeding programs (Gupta *et al.* 2010). Several major genes related to seed morphology have been described in the literature, including the Serpin gene (*Ser5B)*, the speltoid gene (*Q)*, the homoeologous *sphaerococcum* genes (*S*), and the grain protein content gene *Gpc-B1*(Salina *et al.* 2000; Kato *et al.* 2003; Uauy *et al.* 2006; Cane *et al.* 2008). The gene *Ser5B* affects grain characteristics and has been associated with reduced flour yields in Australian wheat germplasm (Rosenkrands *et al.* 1994; Roberts *et al.* 2003). It encodes a grain defense proteinase in the endosperm of developing seeds, possibly playing a role in cell expansion via interaction with storage proteins (Cane *et al.* 2008). The *Q* gene is an important domestication gene because it confers free threshing in wheat. The recessive form of this gene (*q*) is pleiotropic and affects glume shape and tenacity, rachis fragility, spike length, plant height, and spike emergence time. Because of this, cultivated wheat has been selected for the dominant *Q* allele along with several other genes that confer good agronomic type (Sourdille *et al.* 2000). The *Q* allele conferring the free threshing phenotype seen in *T. aestivum* is a dominant gain-of-function mutation from the wild-type *q* allele (Simons *et al.* 2006). It has been mapped to the long arm of chromosome 5A (Kato *et al.* 1999; Kato *et al.* 2003) and cloned (Simons *et al.* 2006). Highly specific markers for rapid characterization of the functional regions of this gene in wheat have been developed (Asakura *et al.* 2009). In newly created (synthetic) hexaploid wheat, *Q* and *Ser5B* are of particular interest for characterization because such germplasm may carry novel alleles of these genes.

Other genes that affect wheat kernel morphology include the homoeologous *S* genes, the gene *Gpc-B1*, and the *rht* dwarfing genes. The *S1*, *S2*, and *S3* genes confer small, spherical grain shape as well as compact spike morphology and were first identified in Indian 'shot wheat' (*T. sphaerococcum*)**.** These

three genes are located on chromosomes 3A (*S3*), 3B (*S2*), and 3D (*S1*) and been mapped near markers *Xgwm2*, *Xgwm566*, and *Xgwm456,* respectively (Maystrenko *et al.* 1998; Salina *et al.* 2000). The *Gpc-B1* gene affects development of the plant during grain-fill period and has pleiotropic effects on seed characteristics (Uauy *et al.* 2006; Waters *et al.* 2009). *Gpc-B1* is located on chromosome 6B and affects grain nutrient content as well as plant senescence through relationships in metabolite translocation (Uauy *et al.* 2006). The *rht* dwarfing genes are well known to be pleiotropic and increase the number of seeds produced per plant (Fischer and Stockman 1986; Keyes 1989; Fischer and Quail 1990; Flintham *et al.*1997; Rebetzke *et al.* 2000; Rebetzke *et al.* 2012). The most widely deployed are *Rht-B1* & *Rht-D1* which decrease sensitivity to endogenous gibberellic acid (gibberellic acid-insensitivity, GAI) and were a critical component of the Green Revolution of the late 1960s (Hedden 2003; Pearce *et al.* 2011). Other *rht* genes which cause reduced stature but are gibberellic acid-responsive (GAR) include *rht-4, rht-5, rht-8, rht-12, rht-13* (Ellis *et al.* 2004; Rebetzke *et al.* 2012). Both GAI and GAR dwarfing genes affect yield components (Rebetzke *et al.* 2000; Rebetzke *et al.* 2012). The *Rht-8* locus on chromosome 2D has pleiotropic effects on grain characteristics, though they are not as pronounced as with other *rht* genes (Flintham *et al.* 1997; Rebetzke *et al.* 2012), and the marker *Xgwm261* has been used to identify QTL for seed characteristics near *Rht-8* (Breseghello and Sorrells 2007). QTL affecting seed size can be found across all chromosomes of wheat, with varying degrees of effect seen for individual QTL (Campbell *et al.* 1999; Dholakia *et al.* 2003; Breseghello *et al.* 2005; Quarrie *et al.* 2005; Huang *et al.* 2006; Sun *et al.* 2009; Gegas *et al.* 2010; Tsilo *et al.* 2010). A recent meta-QTL study has compiled the results of many of these and identified regions on 1A, 1B, 2A, 2D, 3B, 4A, 4B, 4D and 5A that are frequently cited as influencing seed morphology (Zhang *et al.* 2010).

In addition to the genes known to influence kernel characteristics in wheat, several studies describe seed size genes in rice. Two of these are *GW2* and *GS3*. *GW2* confers higher grain weight by impacting the width of rice grains through action of a RING-type E3 ubiquitin ligase (Song *et al.* 2007).

GW2 is located on chromosome two in rice, and has been delimited to a 8.2 kb region flanked by markers *W004* and *W0024* (Song *et al.* 2007). Recently, a wheat ortholog of *GW2* known as *TaGW2* has been cloned (Su *et al.* 2011). *TaGW2* affects seed width and weight and is located on homeologous group 6 in wheat (Su *et al.* 2011). *GS3* affects the length of rice grains via a single base pair change and is located on chromosome three (Guo *et al.* 2009). A marker (*SHJ210*) has been developed which is specific to the functional mutation in *GS3* (Takano-Kai *et al.* 2009). Recent studies have also found QTL on rice chromosome seven that affects both grain length and grain width concurrently (Qiu *et al.* 2012, Shao *et al.* 2012).

A population consisting of doubled-haploid lines from the cross between Synthetic W7984 and Opata M85 was developed as a mapping resource for wheat (Sorrells *et al.* 2011). This population is genetically and phenotypically diverse for several traits including seed morphology. The objective of this study was to identify regions of the genome that are influencing three- dimensional seed shape in the W7984 x Opata M85 and Cayuga x Caledonia (Munkvold *et al.* 2009) populations and compare them to previously identified seed size genes and QTL in wheat and rice.

II. Materials & Methods

Mapping Populations

The primary population used in this study was a subset of 163 lines from the Synthetic W7984X Opata M85 (SynOpDH population) doubled haploid mapping population (Sorrells *et al.* 2011). Opata M85 is a hard red, spring wheat cultivar and Synthetic W7984 is an amphihexaploid wheat developed by crossing 'Altar' durum wheat with *Aegilops tauschii* followed by chromosome doubling. The SynOpDH population was chosen for this study because the parents differ for kernel morphology.

The second population evaluated was the Cayuga x Caledonia doubled haploid (CxC DH) mapping population (Munkvold *et al.* 2009). Two-hundred and eight lines from the CxC DH mapping

population were used for phenotyping. Both parents are winter wheat cultivars with good agronomic qualities, and are adapted to the Northeastern U.S.

Growing Environments and Experimental Design

The SynOpDH population was grown in the greenhouse in the spring of 2008 (GH08) and 2009 (GH09) and at Caldwell field on the Cornell research farm in 2009 (Field09). In the two greenhouse environments, lines were grown in 10cm pots using an augmented design to compensate for seed and space limitations. The population was grown under supplemental lighting with 16h days and 20.0/14.4°C day/night temperature. A subset of the original lines was grown in the greenhouse in 2009; several which were present in 2008 were not included in 2009 due to lack of seed. In the field, lines were grown in single 1.5 meter rows with two replicates using a randomized complete block design.

The CxC DH population was previously grown in two or three field locations per year from 2001 to 2005. Of the 14 original environments in which it was grown, three environments with remnant seed from the 2005 growing season were used for this study. Each location consisted of two replicates of single one-meter rows grown in a randomized complete block design (Munkvold *et al.* 2009). Rows were individually hand-harvested and threshed using a belt thresher. From the original 208 lines a subset of 161 lines were included based on quality of available seed and genotypic data.

Phenotyping

The imaging process for phenotyping was adapted from Breseghello and Sorrells (2007). Twenty-five representative plump, undamaged kernels from each line were photographed. Two photographs were taken of each line including a dorsal view of the kernel, denoted as the horizontal image or 'H image' (Figure 3.1), and a view of the kernel positioned vertically with the embryo end down, denoted the vertical image or 'V image' (Figure 3.2). Following photography, raw images were

processed through the image analysis programs ImageJ (National Institutes of Health, USA, http://rsbweb.nih.gov/ij/) and SHAPE (Hiroshi Iwata, http://lbm.ab.a.u-tokyo.ac.jp/~iwata/shape) to provide individual seed dimensions and elliptic Fourier descriptors of seed shape. Details of the digital image analysis process using these programs are described by Williams *et al.* (2012). Diagrams of the principle component analysis (PCA) scores returned from both the H images and V images of SynOpDH are shown in Figures 3.3 and 3.4, respectively. Diagrams of the PCA scores returned from CxC DH are shown in Figures 3.5 and 3.6. These phenotypic measures are described as HPCs and VPCs, to denote horizontal principle components and vertical principle components, respectively. Thousand kernel weight (TKW) of each line was recorded by weighing all seed from a sample, dividing by total seed number, and multiplying by 1,000. Data were checked for normal distribution and variation for phenotypic measures was tested by one-way ANOVA using JMP8.0 software (SAS Institute Incorporated, www.jmp.com).

Genotyping Known Seed Size Genes & Markers in SynOpDH

Previously reported genes or markers affecting seed characteristics in wheat were screened in the SynOpDH population prior to QTL analysis (Table 3.1). These included a diagnostic cleaved amplified polymorphic sequence (CAPS) marker for the *Ser5B* gene (Cane *et al.* 2008), three CAPS markers specific to regions of the *Q* gene (Asakura *et al.* 2009), a diagnostic simple-sequence repeat (SSR) marker for *Gpc-B1* (Uauy *et al.* 2006, http://maswheat.ucdavis.edu/), SSRs *Xgwm456*, *Xgwm566*, and *Xgwm2* for the *s1, s2, s3* alleles (Salina *et al.* 2000), and SSR *Xgwm261* which was previously reported to be associated with a seed size QTL (Breseghello and Sorrells 2006).

Additionally, markers for rice genes *GW2* and *GS3* were used to genotype the SynOpDH population. The markers *W004* and *W020* were used to genotype SynOpDH lines. The *TaGW2* marker described by Su *et al.* (2011) was not screened because genotyping for this study was performed prior to

publication of their research. The marker *SHJ210* has been used for detecting the functional SNP polymorphism in *GS3* conferring long or short grains (Guo *et al.* 2009). Marker *SHJ210* was used to genotype the SynOpDH population.

DNA was obtained using a seed-extraction protocol described by Kang *et al.* (1998). Following extraction, markers were grouped into annealing temperature categories and were screened for amplification and polymorphism using the parental lines Synthetic W7984 and Opata M85. Polymerase chain reactions were initially set to conditions reported in the literature or the GrainGenes database (Agricultural Research Service, US Department of Agriculture, www.graingenes.org). Those that were not successfully amplified were tested again using a touchdown PCR procedure ending at the recommended annealing temperature. Markers that did not amplify using the touchdown procedure were tested for amplification using gradient PCR spanning annealing temperatures from 48 to 60°C.

Polymorphism was tested using polyacrylamide or agarose gel electrophoresis, depending on individual marker protocols. For SSR and sequence-tagged site (STS) markers, PCR product was run directly on 4% acrylamide gels followed by silver staining. Several of the markers used were CAPS markers (*WSZ1a*, *QR1*, *QR2*, *QR3*, *SHJ210*) and their PCR amplification products were digested using restriction endonucleases as previously described (Rasmussen *et al.* 1996, Asakura *et al.* 2009, Takano-Kai *et al.* 2009). For CAPS markers, PCR products were digested using the recommended restriction enzymes following instructions provided by the manufacturer (NewEngland Biolabs, Ipswich, MA) then run on 3% agarose gels. Markers found to be polymorphic were used for genotyping the entire SynOpDH population.

Linkage Map Construction

SynOpDH

A marker-rich reference map for SynOpDH has been previously generated (Sorrells *et al.* 2011, Poland *et al.* 2012). The map for this population consisted of 1,463 molecular markers, including 114 SSRs and 1,349 diversity array technology (DArT; Diversity Arrays Technology, Yarralumla, Australia) markers covering 2,775.1 cM. A reduced set of markers was selected based on genome coverage with at least one marker every 10cM. A total of 376 markers covering all 21 chromosomes over 2,769.1cM with an average of 7.4cM between markers were used for analysis.

Using the subset of markers from the reference map in conjunction with marker data from known seed size-influencing genes, a reduced map for the SynOpDH population was generated using MapDisto 1.7.0 (Mathias Lorieux, Institut de recherché pour le développement, France, www.mapdisto.free.fr). Marker order was similar to the reference map, but several linkage groups with large distances between markers (>40cM) were broken into smaller groups. This produced a map of 24 linkage groups (Figure 3.7), which was exported for further genetic analysis.

CxC DH

A subset of 161 lines from CxC DH was previously genotyped at 320 loci, including data from 191 SSR, 15 restriction fragment length polymorphism (RFLP), 31 target region amplification polymorphism (TRAP), 72 amplified fragment length polymorphism (AFLP), eight expressed sequence tag-SSR (EST-SSR), and three STS markers. Details of the genotyping and map construction were reported by Munkvold *et al.* (2009). Quantitative trait loci and associated markers from these linkage groups were assigned to wheat chromosomes based on information in the GrainGenes database (Agricultural Research Service, US Department of Agriculture, www.graingenes.org) and their location on the wheat consensus map (Somers *et al.* 2004).

QTL Analyses

SynOpDH

QTL analysis was performed using Win QTL Cartographer v.2.5 (North Carolina State University, www.statgen.ncsu.edu/qtlcart). Map and cross data were imported from MapDisto as Microsoft Excel files. Phenotypic data from each environment was analyzed separately. Composite interval mapping was used for detecting QTL with significance thresholds determined by 1000 permutations at $p<0.01$. Analyses were performed using Model 6, in forward regression, with 10 control markers, and a walk speed of 2.0cM. Single marker regression was used to test significance of markers in the unlinked marker group.

CxC DH

Prior to QTL analysis, all three environments were used to calculate a single best linear unbiased predictor (BLUP) score for each of the kernel morphology phenotypes using JMP8.0 (SAS Institute Incorporated, www.jmp.com). All measures of seed shape were tested for normal distribution and ANOVA was performed using JMP8.0 software.

QTL analysis was performed using QTL Cartographer version 2.5 (North Carolina State University, www.statgen.ncsu.edu/qtlcart/). Traits were analyzed first by single marker regression analysis using all markers to test for linkage groups containing at least one significant locus ($p<0.05$). From this, a reduced version of the map including only linkage groups containing at least one locus significant for any trait from the single marker regression analysis was analyzed using composite interval mapping (CIM). Significance thresholds were set using permutation testing based on 1,000 permutations with significance threshold of $p<0.01$ prior to CIM QTL analysis. QTL Cartographer

parameters were set to CIM Model 6, in forward regression, with 10 control markers, and a walk speed of 2.0cM.

Individual QTL for a trait were defined as the region that the logarithm of odds (LOD) magnitude exceeded the permutation testing threshold. Flanking markers were selected as the two markers closest to either edge of the LOD magnitude once it dropped below the permutation testing threshold. In cases where one of the bordering markers was located close to the highest LOD peak for the QTL, that marker was also designated as the nearest marker. For CxC, individual QTL were described using data for all three environments as BLUP values. For SynOpDH, QTL were described for each of the three environments individually.

III. Results

Screening markers for genes known to affect seed characteristics in the SynOpDH population

Three markers for major genes affecting wheat kernel characteristics were monomorphic in the SynOpDH population. The serpin gene marker *WSZ1a* revealed that both parents, Synthetic W7984 and Opata M85, contain the null allele of *Ser5B* appearing as two DNA fragments of approximately 510 bp and 290 bp for both parents (Figure 3.8). None of the three SNP markers for the speltoid *Q* gene revealed polymorphism between the parents. Similarly, the *Xuhw90* marker for the grain protein *Gpc-B1* locus amplified the same sized PCR product of approximately 126 bp in both parents (Figure 3.9). This indicates that the SynOpDH population does not possess allelic variation for these markers.

Four markers for previously described wheat genes or QTL influencing seed morphology were polymorphic. These included *Xgwm456*, *Xgwm566*, and *Xgwm2* linked to their respective genes *S1*, *S2*, *S3* (Salina *et al.* 2000) and the marker *Xgwm261* associated with QTL affecting seed width on chromosome 2D (Breseghello and Sorrells 2006). Of the four markers, only *Xgwm261* (chromosome 2D) was able to be placed on the reduced SynOpDH map.

For rice *GW2* and *GS3* genes, the markers *W020*, *W004*, and *SHJ210* amplified PCR products in the SynOpDH parents. The *GW2* marker *W020* displayed a complex banding pattern indicating non-specific amplification. The *GW2* marker *W004* amplified a single band in both parents that was monomorphic. In this population, *SHJ210* was able to reliably amplify product but was monomorphic following restriction digests. Furthermore, products of restriction digests showed similar banding patterns between both parents indicative of monomorphism at the functional SNP site of the C165A mutation in rice *GS3* (Takano-Kai *et al.* 2009). This indicates that individuals in this population carry an adenine (A) residue within this gene, characteristic of the C165A allele conferring long grains in rice.

QTL for seed shape in SynOpDH population

All traits (seed length, width, thickness, HPC1-HPC7, VPC1-VPC7, TKW) were normally distributed, and the genotypic source of variance for the traits length, width, thickness, TKW, HPC1, HPC2, HPC3, HPC4, VPC3, and VPC4 were significant (*p* value <0.001) as determined by analysis of variance (ANOVA). VPC1 and VPC2 were discarded since they likely represented error from tilting of the kernel during photography that was not accounted for during image normalization. For more explanation of this, a thorough discussion is provided by Williams *et al.* (2012). Length, width, thickness, TKW, HPC1-HPC4, and VPC3-VPC4 were used in QTL analysis.

Fifty QTL for 10 different traits were differentially detected in three environments in SynOpDH for seed shape characteristics (Table 3.2). Chromosome 5A had the most QTL for shape traits, with 9 QTL for seven different traits describing the seed in the horizontal position as well as VPC4. The QTL with the highest LOD score was for TKW and was located on chromosome 1A with a LOD value of 8.2. The r^2 values ranged from 0.05 for minor QTL to 0.26 for a HPC2 QTL on 1A.

TKW in the SynOpDH population

In the SynOpDH population, seven QTL were detected for TKW. All were detected in the GH08 environment and were located on chromosomes 1A, 4B, 5B (two QTL), 6A, 6D, and 7D. The r^2 values for these QTL ranged from 0.06 to 0.15, with the largest being a QTL detected on 1A, between *Xbarc119* and *Xwmc312*.

Seed Length in the SynOpDH population

Eight QTL were detected for seed length in the SynOpDH population. These were on chromosomes 2A and 6A in the GH09 environment, 2D and 5B in GH08, as well as 4B, 5A, 7A, and 7D in Field09. The r^2 values ranged from 0.06 to 0.15, with the largest occurring on 6A in GH09.

Seed Width in the SynOpDH population

Three QTL were detected for seed width in the SynOpDH population. These were on chromosomes 2A in GH09, 5A in Field09, and 6A in GH08. The r^2 values for these QTL ranged 0.11 to 0.16, with the largest occurring in the GH08 environment on chromosome 6A.

Seed Thickness in the SynOpDH population

For seed thickness a single QTL was detected on chromosome 3B in Field09 between *wPt-8079* and *wPt-9066*. The nearest marker was *wPt-8079* and the QTL had an r^2 value of 0.09.

Shape Descriptor HPC1 in the SynOpDH population

In the SynOpDH population, HPC1 detected uniform widening of the kernel along its length (Figure 3.3). Seven QTL were detected on chromosomes 1A (GH09), 1B (GH08), 2A (GH09), 2B (GH08), 2D (nearest *Xwmc112* in three environments GH08, GH09, Field09), 4B (GH08), and 5A (GH09). The r^2

values ranged from 0.05 to 0.16, with the highest value belonging to a QTL detected on chromosome 4B near marker *Xwmc710*. Note that HPC1 QTL associated with *Xwmc112* on 2D were detected across all environments. Additionally, the QTL on chromosome 2B is near the predicted position of *Rht-B1*, which would be located between *Xwmc710* and *Xmc511* on the SynOpDH map based on comparison with the wheat composite and consensus maps (Somers *et al.* 2004, Agricultural Research Service, US Department of Agriculture, www.graingenes.org).

Shape Descriptor HPC2 in the SynOpDH population

In the SynOpDH population, HPC2 detected widening of the kernel at either lateral end (Figure 3.3). Eight QTL were detected on chromosomes 1A (nearest *wPt-8644* in two environments GH08, Field09), 1D (GH09), 2D (Field09), 3B (two QTL, both GH08), 5A (two QTL, both GH09), and 7A (GH09). The r^2 values for these QTL ranged from 0.05 to 0.26 for QTL near *wPt-8644* on 1A.

Shape Descriptor HPC3 in the SynOpDH population

In the SynOpDH population, HPC3 detected the variation in a rectangular or evenly ovoid shape of the wheat kernel. Five QTL were detected on chromosomes 1B (Field09), 2B (two different QTL, both in GH08), 2D (GH09), and 5A (GH09). QTL for this trait were environment-specific, none being detected in multiple environments. The r^2 values ranged 0.08 to 0.22, with the largest r^2 detected on 2B near *Xbarc18*. The QTL near *Xbarc18* does not co-locate with the predicted position of *PpdB1*. Comparison to the composite and consensus maps (Somers *et al.* 2004, Agricultural Research Service, US Department of Agriculture, www.graingenes.org) predicts that *PpdB1* would be expected to occur near the end of the short arm of chromosome 2B; notably the QTL near *Xbarc18* is more proximally located within the linkage group. However, *Xbarc18* is close to near the predicted position of *Rht-B1*.

Shape Descriptor HPC4 in the SynOpDH population

In the SynOpDH population, HPC4 detected variation in the taper of a wheat kernel at either end. Eight QTL were detected on 2B (three different QTL, all in GH08), 2D (two different QTL in Caldwell 09 and GH09), 3A (GH08), and two on 5A (same QTL in GH09 and Field09, independent QTL in GH08). Furthermore, the QTL detected on 5A was significant in multiple environments. The r^2 values ranged from 0.08 to 0.24. The QTL near marker *wPt-2214* on chromosome 2B had the highest r^2 value.

Shape Descriptors VPC3, VPC4 in the SynOpDH population

In the SynOpDH population, VPC3 detected flattening or how 'heart-shaped' a kernel was when viewed on end (Figure 3.4). VPC3 could also be interpreted as a measure of thickness. In the SynOpDH population, VPC4 detected slight changes in how circular the perimeter of a wheat kernel was when viewed on end. Only three significant QTL were detected using shape descriptors VPC3 or VPC4. One QTL for VPC3 on 1A and two QTL for VPC4 on 4A and 5A were detected. The r^2 values for these QTL ranged from 0.1 for the QTL detected by VPC3 on chromosome 1A to a 0.22 for the QTL for VPC4 detected on 5A.

QTL for seed shape in the CxC DH population

One-way ANOVA revealed that experimental lines in the CxC DH population have significant differences for all phenotypic traits measured using ImageJ (p-values <0.001). Using SHAPE, one-way ANOVA found significant differences in all phenotypic traits as measured by EFD measures of seed shape (p-values <0.001) except for VPC1 and VPC2. VPC1 and VPC2 likely represented error due to tilting of the kernel during photography not accounted for by image normalization. For more explanation of this, a thorough discussion is provided by Williams *et al.* (2012).

Using BLUPs calculated from the three environments sampled, 32 QTL were detected for the multiple seed morphology traits (Table 3.3). A number of QTL for different seed morphology traits co-localized. Proportion of variance explained by each of these ranged from 7%-21% and the average r^2 was 0.11. One QTL was detected on chromosome 1A, two on 1D, five on 2A, three on 2B, one on 2D, four on 3B, three on 3D, one on 4B, five on 4D near the *rht* locus, five on 6D, and two on 7D. LOD scores for these QTL ranged from 3.7 to 18, with an average of 6.96 (Table 3.3).

TKW in the CxC DH population

Three QTL were detected on 4D, 6D, and 7D. The r^2 values for these ranged from 0.09 – 0.14. The largest value was for the QTL on 4D near *RHT-DF-MR2*, which explained 14% of the variability.

Seed Length in the CxC DH population

Six QTL were detected for seed length on chromosomes 1D, 2A (two different QTL), 3B, 3D, and 6D. The r^2 values ranged from 0.08 to 0.21. The QTL on 3D had the largest value of r^2, which was 0.21, and the nearest marker was *Xbarc1161*.

Seed Width in the CxC DH population

Five QTL were detected on chromosomes 2A, 2B, 4B, 4D, and 6D. The r^2 values ranged from 0.07 to 0.19. The QTL on 4D nearest marker *RHT-DF-MR2* had the largest r^2 value of 0.19.

Seed Thickness in the CxC DH population

Four QTL were detected on chromosomes 2A, 3B, 4D, and 7D. The r^2 values ranged from 0.08 to 0.16. The largest QTL was on chromosome 2A nearest *Xgwm339* and *Xgwm515*, and had an r^2 value of 0.16.

Shape Descriptor HPC1 in the CxC DH population

In the CxC DH population, HPC1 detected the concurrent widening and shortening of the kernel, similar to length to width ratio (Figure 3.5). Seven QTL were detected on 1D, 2A, 2B, 3B, 3D 4D, and 6D. The r^2 values for these QTL ranged from 0.07 to 0.18, with the highest values belonging to QTL on 1D near *Xbarc0346* and on 3D near *Xbarc1161*.

Shape Descriptor HPC2 in the CxC DH population

In the CxC DH population, HPC2 detected tapering at either end of the wheat kernel (Figure 3.5). Only one QTL was detected on chromosome 2D near *Xbarc1123* ($r^2 = 0.14$).

Shape Descriptor HPC3 in the CxC DH population

In the CxC DH population, HPC3 detected simultaneous tapering of the kernel at both ends, describing how uniformly rounded a kernel was (Figure 3.5). Two QTL were detected on chromosomes 3B and 3D. The r^2 values for these were 0.16 and 0.08, respectively. The nearest marker to the QTL on 3B was *E34M49161L*. The nearest marker to the QTL on 3D was *Xbarc1161*.

Shape Descriptor HPC4, HPC5 in the CxC DH population

In the CxC DH population, HPC4 & HPC5 did not detect any significant QTL.

Shape Descriptor VPC1, VPC2 in the CxC DH population

In the CxC DH population, one QTL was detected for VPC1 on chromosome 4D near *Xbarc1118*, although the validity of this QTL is questionable until the issue of shape normalization is resolved. (See comments in first paragraph of this section.) No QTL were detected for VPC2.

Shape Descriptor VPC3 in the CxC DH population

In the CxC DH population, VPC3 described crease depth and increasing kernel thickness (Figure 3.6). Two QTL were detected for this trait, one on chromosome 2B near *wPt-2430* and *Xwmc360*, and one on chromosome 1A near *CFA2129_RTL*. These QTL had r^2 values of 0.08 and 0.07, respectively.

Shape Descriptor VPC4 in the CxC DH population

In the CxC DH population, VPC4 also described crease depth, but without the concurrent changes to thickness that were detected by VPC3 (Figure 3.6). One QTL was detected for VPC4 on chromosome 6D. The r^2 for this QTL was 0.20, and the nearest marker was *Xcfd37*.

IV. Discussion

Several QTL were detected near predicted positions of known genes evaluated in this study. The *Ser5B* gene, *Q* gene, and *Gpc-B1* markers were monomorphic in SynOpDH, preventing placement on the linkage map. However, the genomic locations of these genes are well characterized in other populations and lack of polymorphism suggests that phenotypic variation in SynOpDH is not associated with markers previously reported to be linked to these genes (Kato *et al.* 2003; Uauy *et al.* 2006; Cane *et al.* 2008). SynOpDH was polymorphic for markers linked to the *S* genes on homoeologous group three and several QTL were discovered on chromosomes 3A and 3B (Table 3.2). Unfortunately the markers for *S1, S2, and S3* could not be placed on the linkage map. However in this study, the markers were not associated with shape traits based on single marker analysis. A QTL for HPC2 was located near marker *wPt7152* that corresponded to the predicted position of *S2* (between *Xwmc43* and *Xwmc418* on 3B) based on comparison to the consensus map (Somers *et al.* 2004). The predicted position of *S3* did not

correspond to the QTL detected on 3A in SynOpDH based on comparison to the Nanda2419 x Wangshuibai map (Xue *et al.* 2008). No QTL were detected on 3D in the SynOpDH population near the predicted position of *S1*.

In the SynOpDH population, significant QTL for seed length, width, HPC1 through HPC4, and VPC4 were located on chromosome 5A, though no QTL were detected on chromosome 5A in CxC DH. Since the *Q* gene is on 5A, knowing its chromosome location can aid in determining if a *Q* allele is contributing to 5A QTL. Screening of all known SNP polymorphisms across the *Q* gene in SynOpDH did not detect any difference between parental lines, suggesting that W7984 and Opata M85 have the dominant form of the *Q* allele and another gene on 5A may be affecting kernel morphology. Due to lack of polymorphism for described PCR-based *Q* markers and difficulty in estimating position of *Q* in the SynOpDH population based on previous RFLP-based mapping reports, further assessment of the region harboring *Q* using nearby, polymorphic markers is needed. The Serpin gene (*Ser5B*) on 5B is predicted to be on the long arm of chromosome 5B based on previous reports (Cane *et al.* 2008), though there is little information on specific markers surrounding the gene. In SynOpDH two QTL on 5B for seed length (near *wPt5688*) and TKW (near *wPt5688*, *Xbarc59*) were detected in the GH08 environment. The two QTL were located at opposite ends of 5B (Figure 3.7), and could potentially be related to *Ser5B* given lack of detailed marker information surrounding the gene. However, the lack of variation for the genic marker *WSZ1a* used to characterize the *Ser5B* locus in SynOpDH suggests that observed 5B QTL were unlikely to be associated with the *Ser5B* gene. The predicted location of *Gpc-B1* would be near the centromere of chromosome 6B (Distelfield *et al.* 2006). No QTL were detected on 6B in SynOpDH and markers for the *Gpc-B1* locus were monomorphic. In the case of *Gpc-B1*, lack of molecular variation described in this study fits well with casual observations of a lack of stay-green phenotypes normally associated with *Gpc-B1* (Uauy *et al.* 2006). In the CxC DH population, no QTL were detected for kernel

morphological characteristics on chromosomes 5B or 6B. Therefore, differences in seed size in SynOpDH or CxC DH are unlikely to be the result of allelic variation for *Ser5B* or *Gpc-B1*.

It was difficult to compare QTL in the SynOpDH population to the location of known rice seed size genes, because no variation for the *GS3* or *GW2* markers was observed. QTL were detected in SynOpDH for seed length and width on chromosome 5AL near *wPt-8262* (Table 3.2), potentially corresponding to *GS3* since rice chromosome three has been comparatively mapped to regions of 5A/5B/5D in wheat (Wilson *et al.* 1999, Sorrells *et al.* 2003). The *GS3* gene is located on the short arm of rice chromosome three. This is particularly suggestive of an orthologous gene given results detecting associations between marker *Xwmc150a* on 5A and seed length in a diverse collection of wheat germplasm (Breseghello and Sorrells 2006) and that a seed length QTL was detected on 5AL in the SynOpDH population. However, screening for the putative *GS3* ortholog in wheat using *SHJ210* in the SynOpDH population revealed no polymorphism. The marker *SHJ210* detects the functional SNP polymorphism of *GS3*, producing a mutant 'A allele' conferring long grains or wild type 'C allele' for shorter grains (Takano-Kai *et al.* 2009). Because both seed length and width QTL were detected on chromosome 5AL of SynOpDH, the gene underlying this QTL may be an allelic variant of *GS3* (or a novel gene) which also affects width or thickness of seeds. Furthermore, the most significant QTL detected on 5A was for VPC4, which was also near the region identified in the SynOpDH population affecting seed length and width.

In rice, the *GW2* gene influences grain width and is located on rice chromosome two, which has been comparatively mapped to homoeologous group six in wheat (Agricultural Research Service, US Department of Agriculture, www.graingenes.org). The STS marker *W004* developed for *GW2* (Song *et al.* 2007) did not detect polymorphism between the parental lines in this study. Subsequent to the genotyping performed for this study, Su *et al.* (2011) have reported cloning of a wheat ortholog of *GW2* which is located on the homeologous group 6 chromosomes. Although the rice marker *W004* used in

this study was observed to be monomorphic, further screening of SynOpDH using the markers developed by Su *et al.* (2011) for *GW2* is warranted, since phenotypic effects of QTL for width, length, and TKW on 6A in SynOpDH may be related to the function of the *GW2* ortholog (width).

There is interest in seed shape QTL in wheat because kernel morphology may impact quality traits such as milling yield. Flour yield data was not available for the SynOpDH population. However, we compared QTL in SynOpDH to previously described QTL for seed shape and flour yield to QTL in CxC DH (Williams *et al.* 2012). In the CxC DH population, QTL co-localized for flour yield and seed width on chromosome 2B, flour yield and seed length on 1D, and flour yield and seed thickness on 7D. Additionally the SynOpDH results in this study can be compared to a previous report that mapped quality traits in the International *Triticeae* Mapping Initiative (ITMI) population (Nelson *et al.* 2006). The ITMI population is a recombinant inbred line (RIL) population constructed using the same parents as the SynOpDH population. While Nelson *et al.* (2006) did not phenotype seed morphology, they did focus on a number of grain quality characteristics including flour yield. In their report, QTL on 5D, 4A, and 4D affected flour yield. In this study no QTL were found on 5D or 4D, though a QTL for VPC4 was found on 4A. Comparison to the wheat consensus map shows the 4A QTL in SynOpDH to be near the end of the linkage group (~25cM) whereas the 4A QTL reported by Nelson *et al.* (2006) is more proximally located (~80cM). The SynOpDH population seems to be a promising source of alleles from 1A, 3A, and 3B to select grain ideotypes having a larger internal volume to surface area ratio that may contribute to increased flour yield.

Loci on homoeologous group two may harbor valuable alleles for simultaneously impacting HPC descriptors of seed shape and quality parameters of a cultivar. On 2B, direct measures of kernel width (CxC DH) and uniform widening along the length of the kernel (SynOpDH; HPC1, HPC3, HPC4) may be conditioned by the same locus contributing to flour yield or grain protein content. QTL for HPC1, HPC3, and HPC4 were detected on 2B in the SynOpDH population near a region influencing width and flour

yield in CxC DH (Williams *et al.* 2012). On 2D near the *Ppd* locus, QTL were detected in SynOpDH affecting all HPC measures. In the ITMI population, QTL were detected in a similar region that impacted grain protein content (Nelson *et al.* 2006). This region on 2D contains genes for *Ppd* (photoperiod response) and *Tg1* (glume tenacity). Studies of kernel growth have shown that light penetration into the floral cavity and glume tenacity can influence grain morphology as well as pericarp thickness (Millet and Pinthus 1984; Raju and Srinivas 1991). In SynOpDH there is evidence of segregation for glume tenacity among lines. Whether the shape QTL observed on 2D result from light sensing effects of *Ppd* or effects of *Tg1* on glume tenacity is an interesting question. Furthermore, genes affecting kernel morphology on homeologous group 2 in wheat may be orthologous to those underpinning recently described rice chromosome 7 QTL that affect multiple dimensions of the seed, based on comparative mapping that has found similarity between rice chromosome 7 and segments of chromosomes 2B and 2D of wheat (Qiu *et al.* 2012; Shao *et al.* 2012; Agricultural Research Service, US Department of Agriculture, www.graingenes.org).

QTL on 7D impact TKW, thickness, and flour yield, while QTL on 7A impact length, HPC2, and grain protein though effects on individual dimensions of the kernel appear to be population specific. On chromosome 7D, QTL for TKW (*Xcfd21*) in the SynOpDH population corresponded to thickness/TKW/flour yield QTL in CxC DH (*Xwmc150b*, Williams *et al.* 2012). Adjacent to this locus a QTL for length was also identified in SynOpDH. In the ITMI population QTL for grain protein were detected on 7AS (Nelson *et al.* 2006), and in the SynOpDH population, QTL for length and HPC2 were detected on 7AL. However, owing to variable allelic effects on phenotype between the populations compared, quality assessments need to be made directly on SynOpDH to confirm relationship of the QTL to both seed shape and quality parameters. Alternately the QTL regions could be introgressed into another population or isolated in near isogenic lines and then phenotyped to the determine impact of these loci.

QTL identified in SynOpDH on 3A and 3B could be used in conjunction with QTL from 1A to select seeds with greater internal volume to surface area ratios. Interestingly, the QTL on 3A and 3B in SynOpDH both seem to affect uniform widening along the length of the wheat kernel (HPC1) and thickness, and explain more phenotypic variation for these traits than other QTL. These loci could be used with the QTL for HPC2 on chromosome 1A (near *wPT-8644*) to select for more rectangular seed shapes with greater internal volume. The use of genotypic data to select for these QTL in wheat breeding may overcome the previous obstacle of yield component compensation, but only if other yield components are mapped as well. Ability to select for seeds with greater internal volume to surface area may improve flour yield as postulated by geometric models (Marshall *et al.* 1983), though the effectiveness of the selection technique would need to be demonstrated empirically.

The most interesting QTL detected in this study were those on chromosomes 1A, 2D, and 5A. These QTL affected seed shape across multiple environments or affected multiple dimensions of the seed. They provide validation of seed size and shape QTL identified in other populations (Breseghello and Sorrells 2006; Huang *et al.* 2006; Gegas *et al.* 2010; Zhang *et al.* 2010). These QTL also demonstrate the utility of using all three spatial dimensions as a way to distinguish the impact of individual loci on a complex morphological phenotype.

In the SynOpDH population, environmentally stable QTL for EFDs were detected on 1A (Figure 3.10) that corresponded to seed size QTL (typically as TKW) reported in other studies (Gegas *et al.* 2010; Zhang *et al.* 2010). Chromosome 1B has been frequently cited as harboring TKW QTL, though significant regions have been detected on 1A and 1D as well (Zhang *et al.* 2010). Similar to the QTL affecting HPC1 and HPC2 in the SynOpDH population, QTL on 1A have been described for seed length to width relationships (Gegas *et al.* 2010). Inclusion of both direct measures of seed dimensions (length, width, thickness) as well as EFD shape descriptors (HPCs, VPCs) for comparison allow the unique identification of QTL affecting complex seed shape phenotypes. In the SynOpDH population, several markers including

wPt-5660, *wPt-8644*, *wPt-1709* identified QTL for HPC1, HPC2, and VPC3 on chromosome 1A. Rather than determining maximum width or length *per se*, this locus contributes to uniformity of seed width along its length as indicated by the lack of QTL for direct measures of width or length in the same region. These results support the findings of Gegas *et al.* (2010), who described QTL on 1A affecting the length to width relationship of the seed. If HPC1 or HPC2 are related to flour yield and prove to be more stable than traditionally used seed size measurements such as TKW, this region and its associated markers may be useful to breeding efforts. Notably, Breseghello and Sorrells (2006) reported a flour yield QTL on 1A. In CxC DH, a QTL for VPC3 (related to crease depth) was found on 1A. While not directly measured in this study, the role of the crease in influencing seed characteristics remains interesting because surrounding vasculature is involved in nutrient transport to developing grains (Cochrane 1983; Ugalde and Jenner 1990a; Ugalde and Jenner 1990b). Though the vasculature does not appear to limit the rate or duration of the grain filling period, it does establish the developmental patterning of the grain early on and could therefore impact morphology (Lingle and Chevalier 1983, Drea *et al.* 2005). Several previous reports of QTL influencing flour yield did not report QTL at this chromosome location (Campbell *et al.* 2001; Kuchel *et al.* 2006; Lehmensiek *et al.* 2006; Nelson *et al.* 2006; Smith *et al.* 2011; Tsilo *et al.* 2011; Wang *et al.* 2012).

QTL detected on 2D affected HPC1 and may identify environmentally stable pleiotropic effects of *Ppd* or *Tg1* on seed shape or development. In the SynOpDH population, the region surrounding the photoperiod response gene *Ppd* on 2D had multiple QTL detected that corresponds approximately to the position of previously reported seed morphology meta-QTL (MQTL15, MQTL16; Zhang 2010). In another wheat mapping population, Reed X Grandin, the markers *Xwmc18* and *Xgwm30* were able to identify QTL affecting seed width on 2D, although they are far from *Ppd* (Breseghello and Sorrells 2006). Similarly a flour yield QTL was reported in a region different than *Ppd* by Wang *et al.* (2012). A wider sampling of allelic diversity from an association mapping study of kernel morphology found *Xwmc111*,

near *Ppd*, to be associated with changes in seed width (Breseghello and Sorrells 2006). A recent study described flour yield QTL from a similar region of 2D (Smith *et al.* 2011). Alleles of *Ppd* affect inflorescence growth and relative growth rate of the floral apex (Scarth 1985). In SynOpDH no QTL for seed width were detected on 2D near *Ppd*. However, QTL were detected near *Ppd* across all three environments for HPC1, which generally describes the uniform widening along the length of the kernel. *Ppd* may be a principle cause of this variation due to pleiotropy, but often the confidence intervals of QTL did not encompass the *Ppd-1* marker (Figure 3.11). This suggests further fine-mapping could reveal a gene other than *Ppd* affecting seed characteristics. The marker *Xwmc112* was located directly beneath the environmentally stable QTL on 2D for HPC1. Previously *Xwmc112* was used to characterize the *Tg1* gene affecting glume tenacity in the ITMI population (Nalam *et al.* 2007). Casual observation of variation in threshing characteristics among SynOpDH lines, with tough threshing coming from Synthetic W7984, was initially thought to be conferred by the *Q* locus but genotyping determined that both parents carry the dominant form of *Q*. Thus, variation in threshing characteristics was not due to pleiotropic effects of *Q* but rather another gene, possibly *Tg1*, underlying stable QTL for HPC1 detected by *Xwmc112* in the SynOpDH population. Furthermore, given the variation in threshing characteristics observed in this population, the QTL affecting uniform seed widening (HPC1) in SynOpDH could be a pleiotropic effect of *Tg1*. Several earlier studies (Engledow 1920; Lamba 1949; Radley 1981; Millet 1986) have reported that glume length and width (though not depth of the floral cavity) are correlated to seed shape. Light penetration into the floral cavity may play a significant role in regulating grain growth as well (Millet and Pinthus 1984). In rice, loosening of the glumes has been found to affect development, shape, and surface roughness of the rice grain (Raju and Srinivas 1991). Additionally, glumes affect thickness of the pericarp in both rice and wheat (Millet and Pinthus 1984; Raju and Srinivas 1991). The physiological relationship between the adhesion of the glumes and pericarp may potentially constrain the development of the seed after length is determined, although this requires verification by further

experiments. More practically these results indicate allelic variation for seed shape exists due to whatever gene underlies QTL near *Xwmc112*, though phenotypic effects need to be assessed in the target population following introgression. Successful manipulation of this region will depend on use of genotypic data during selection since phenotypic evaluation alone may select different loci on 2D affecting seed width, or perhaps seed length if a derived measure of seed size such as TKW is used.

QTL on 5A were associated with multiple measures of seed shape, and specifically identify a phenotype which could be used to select a desirable, environmentally stable locus. When examining both greenhouse and field environments, QTL for HPC4 were detected on 5A. When examining only the field environment, QTL for VPC4, seed length, and thickness were detected in addition to HPC4 on 5A. QTL affecting seed shape and TKW on chromosome 5A have been reported previously, and meta-QTL analysis has placed clusters of seed shape QTL on this chromosome (Kato *et al.* 2000; Gegas *et al.* 2010; Zhang *et al.* 2010). Additionally, QTL on 5A were found to be associated with length (Breseghello and Sorrells 2006), grain yield, and test weight (Huang *et al.* 2006). These previously reported QTL were located on opposite ends of the chromosome, with those reported by Huang *et al.* (2006) occurring around *Xcfd39* on the long arm of 5A, and the associations reported by Breseghello and Sorrells (2006) surrounding *Xwmc150* on the short arm. In the SynOpDH population the QTL affecting HPC4, VPC4, seed length, and width were detected by marker *wPt-8262* near the location of QTL for test weight and grain yield reported by Huang *et al.* (2006). The region described by Breseghello and Sorrells (2006) affecting length is located in a different region of the linkage group, with a predicted position closer to HPC1 and HPC2 QTL detected on the short arm. Phenotyping HPC4 may allow selection of a locus bordered by *Xgwm304* and *wPt-8226* on 5AS that impacts length and width across multiple environments based on observation of differential detection of shape QTL based on environment.

When comparing results of QTL analyses between SynOpDH and CxC DH, some overarching patterns emerge. These include different allelic contributions from the constituent genomes (A,B,D) and

disparate effects of specific loci on seed shape. The differing effects of seed shape loci between populations have potential implications for related quality traits. These observations raise new questions concerning the use of functional molecular variation and the phenotyping process in plant breeding.

As reported here, useful allelic contributions from the constituent genomes may vary based on whether populations are derived from adapted materials or from exotic germplasm. Comparing between the two populations, SynOpDH only had 10 of 50 (~20%) QTL detected in the D genome, whereas CxC DH had 18 of 32 (~56%) QTL detected in the D genome. Discrepancies between the populations indicate that the genetic factors contributing to seed shape may vary uniquely based on contributions from each of the three genomes in hexaploid wheat.

Based on these results, the effects of seed morphology QTL are often population specific and may require several measures of shape to identify QTL affecting complex morphological phenotypes. For example, QTL on 2B in the CxC DH population affect width and flour yield whereas in the SynOpDH population they affect HPCs. Effects of these alleles from the SynOpDH population may contribute to subtle changes in overall seed shape, such as more uniform lateral widening, potentially improving quality characteristics like flour yield. Loci described on 1A, 3A, and 3B in the SynOpDH population may be particularly promising in this respect. Seed morphology alleles contributed by the A/B genome of synthetic wheat could have novel effects once integrated into adapted germplasm. Furthermore, alleles with subtle effects on complex shape phenotypes may go undetected or be improperly characterized due to limitations of phenotyping using only dimensions of major axes.

Evaluation of grain shape in the SynOpDH population revealed numerous QTL affecting the complex dimensions of seeds. These were unique in comparison to those detected in CxC DH, though loci on 1A, 2D, and 5A in the SynOpDH population are potential targets for breeding and validate previously described seed morphology QTL. The regions on chromosomes 1A and 2D are useful because

they impact rectangular shape of grains and width (respectively) and were environmentally stable. The multiple QTL on chromosome 5A can be used to select for either grain width or thickness. The locus affecting thickness on 5A is promising as a selection target because use of the phenotype HPC4 may allow selection for desirable seed shape characteristics across multiple environments.

Table 3.1. List of markers screened in SynOpDH

Marker	Associated Gene	Successful Amplification	Polymorphic	Reference
SHJ210	rice *GS3*	yes	no	Takano-Kai *et al.* 2009
W004	rice *GW2*	yes	no	Song *et al.* 2007
W020	rice *GW2*	yes	non-usable	Song *et al.* 2007
Xgwm261	seed QTL/*rht-8*	yes	yes	Breseghello and Sorrells 2006
Xuhw89	*GpcB1*	yes	no	Uauy *et al.* 2006
QR1	*Q* region 1	yes	no	Asakura *et al.* 2009
QR2	*Q* region 2	yes	no	Asakura *et al.* 2009
QR3	*Q* region 3	yes	no	Asakura *et al.* 2009
Xgwm456	*S1*	yes	yes	Salina *et al.* 2000
Xgwm566	*S2*	yes	yes	Salina *et al.* 2000
Xgwm2	*S3*	yes	yes	Salina *et al.* 2000
WSZ1a	*Ser5B*	yes	no	Rosenkrands *et al.* 1994

Table 3.2. QTL detected for seed shape characteristics in SynOpDH

Trait	Environment	Chromosome	Bordering Markers	Nearest Marker	LOD	r^2
Length	GH09	2A	*wPt-9320, Xgwm294*	*wPt-5351*	5.2	0.10
	GH08	2D	*wPt-7466, Xwmc167*	*wPt-7921*	5.0	0.13
	Field09	4B	*Xwmc710, wPt-3991*	*wPt-3991*	4.0	0.07
	Field09	5A	*Xcfa2141, wPt-5231*	*wPt-8262*	5.3	0.09
	GH08	5B	*wPt-6348, wPt-5514*	*wPt-5688*	6.1	0.14
	GH09	6A	*wPt-0139, Xwmc570*	*wPt-0902*	6.6	0.15
	Field09	7A	*wPt-0321, Xbarc121*	*wPt-0288*	4.4	0.06
	Field09	7D	*Xcfd21, wPt-0231*	*wPt-7278*	6.4	0.14
Width	GH09	2A	*Xgwm294, wPt-3136*	*wPt-3136*	4.6	0.11
	Field09	5A	*Xwmc415, wPt-8262*	*Xcfa2141*	4.8	0.13
	GH08	6A	*wPt-0902, Xwmc570*	*Xwmc807*	4.8	0.16
Thickness	Field09	3B	*wPt-8079, wPt-9066*	*wPt-8079*	3.9	0.09
HPC1	GH09	1A	*wPt-4658, wPt-5660*	*wPt-5660*	4	0.07
	GH08	1B	*wmc694,wPt-1818*	*wPt-8280*	5	0.12
	GH09	2A	*wPt-5251,wPt-9320*	*wPt-9320*	4	0.07
	GH08	2B	*wPt-6174, wPt-3755*	*Xbarc18*	6	0.15
	Field09	2D	*Xcfd56, Ppd-1*	*Xwmc112*	5.6	0.1
	GH08	2D	*wPt-9070,wPt-4406*	*Xwmc112*	5	0.13
	GH09	2D	*wPt-9070, Ppd-1*	*Xwmc112*	7	0.13
	GH08	4B	*wPt-9393,Xwmc511*	*Xwmc710*	6	0.16
	GH09	5A	*wPt-2768,wPt-8226*	*wPt-8226*	4	0.05
HPC2	Field09	1A	*wPt-7972,Xbarc17*	*wPt-8644*	3.5	0.1
	GH08	1A	*wPt-7972,Xbarc17*	*wPt-8644*	7	0.26
	GH09	1D	*wPt-8454,wPt-1006*	*wPt-1006*	3.3	0.05
	Field09	2D	*wmc112,wPt-5222*	*Ppd1*	3.5	0.1
	GH08	3B	*wPt-0250,Xwmc43*	*wPt-0250*	3.2	0.11
	GH08	3B	*wPt-1366,wPt-4364*	*wPt-7152*	4	0.12
	GH09	5A	*wPt-7381,Xgwm304*	*wPt-7381*	4.4	0.1
	GH09	5A	*Xgwm304,wPt-8226*	*wPt-2768*	4.4	0.1
	GH09	7A	*wPt-9314,wPt-0321*	*Xcfa2028*	4.5	0.12
HPC3	Field09	1B	*wPt-1818,wPt-3457*	*wPt-5061*	3	0.08
	GH08	2B	*wPt-7619,Xbarc18*	*wPt-1064,Xbarc18*	4.8	0.22
	GH08	2B	*wPt-8916,wPt-7408*	*wPt-2214*	4.8	0.2
	GH09	2D	*Xcfd56, wPt-9070*	*wPt-9070*	3.3	0.08
	GH09	5A	*Xcfa2141,wPt-5231*	*wPt-5231*	4.8	0.11
HPC4	GH08	2B	*wPt-1064,wPt-6174*	*wPt-1064*	3	0.12
	GH08	2B	*wmc474, wPt03755*	*Xbarc18*	4.2	0.12
	GH08	2B	*wPt-8916,wPt-7408*	*wPt-2214*	7	0.24
	Field09	2D	*Xcfd56,wPt-4406*	*wPt-9070*	3.8	0.11
	GH09	2D	*Xwmc112,wPt-4406*	*wPt-4406*	4	0.11
	GH08	3A	*wPt-1002,tPt-6234*	*tPt-6376*	5	0.2
	GH08	5A	*wmc415,wPt-8262*	*Xcfa2141*	5	0.19
	Field09	5A	*Xcfa2141,wPt-5231*	*wPt-8262*	3.4	0.09
	GH09	5A	*wPt-8262,wPt-5231*	*wPt-8262*	3.3	0.08
VPC3	Field09	1A	*Xbarc119,wPt-1709*	*wPt-1709*	3.2	0.1
VPC4	GH09	4A	*wPt-7391,Xwmc650*	*wPt-8019*	3.3	0.11
	Field09	5A	*Xcfa2141,wPt-5231*	*wPt-8262*	6.8	0.22
Thousand Kernel Weight	GH08	1A	*Xbarc119, Xwmc312*	*wPt-1709*	8.2	0.14
	GH08	4B	*wPt-8555, Xwmc511*	*wPt-9393*	6.1	0.11
	GH08	5B	*wPt-5688, wPt-0929*	*wPt-5688*	4.2	0.07
	GH08	5B	*wPt-6880, wPt-3115*	*Xbarc59*	6.1	0.10
	GH08	6A	*wPt-0139, Xwmc570*	*Xwmc807*	8.1	0.15
	GH08	6D	*wPt-5331, Xcfd60*	*wPt-9589*	4.0	0.06
	GH08	7D	*csLV34, wPt-7278*	*Xcfd21*	5.0	0.12

Table 3.3. QTL detected for seed shape characteristics in CxC DH

Trait	Chromosome	Bordering Markers	Nearest Marker	LOD	r^2
Length	1D	*Xwmc264, Xbarc0045*	*Xbarc0346*	6.3	0.08
	2A	*E40M59257Y, Xgwm256b*	*E40M59257Y*	4.8	0.08
	2A	*Xgwm639b, E36M602270L*	*Xwmc150a*	5.5	0.10
	3B	*Xbarc0077, E35M49161L*	*Xbarc0077*	7.7	0.16
	3D	*Xcfd70, Xgwm645*	*Xbarc1161*	18.0	0.21
	6D	*Xbarc0204, Xbarc0096*	*Xbarc0204*	8.0	0.10
Width	2A	*Xgwm339, Xwmc522*	*Xgwm339*	3.7	0.09
	2B	*Xbarc0328, Xwmc360*	*wPt-2430*	4.2	0.08
	4B	*Xwmc435,Xbarc0163*	*Xgwm192*	5.0	0.07
	4D	*Xbarc0217, Xbarc1118*	*RHT-DF-MR2*	8.0	0.19
	6D	*Xbarc0096, Xcfd01*	*Xbarc0196*	4.4	0.07
Thickness	2A	*Xgwm339, Xgwm095*	*Xgwm339, Xgwm515*	5.9	0.16
	3B	*E35M49161L, TRAP_telos12-32*	*Xbarc0229*	5.4	0.08
	4D	*Xbarc0217, Xbarc1118*	*RHT-DF-MR2*	5.0	0.10
	7D	*Xwmc150b, Xbarc0126*	*Xbarc0172, Xgwm473*	4.6	0.09
HPC1	1D	*Xwmc264, Xbarc0045*	*Xbarc0346*	16.0	0.18
	2A	*Xgwm639b, E36M602270L*	*Xwmc150a*	6.0	0.07
	2B	*Xwmc770, Xbarc0328*	*Xgwm429*	7.0	0.07
	3B	*Xbarc0077, E35M49161L*	*Xbarc0077*	11.0	0.09
	3D	*Xcfd70, Xgwm645*	*Xbarc1161*	10.8	0.18
	4D	*Xbarc0217, Xbarc1118*	*RHT-DF-MR2*	11.0	0.09
	6D	*Xbarc0204, Xbarc0096*	*Xbarc0204*	5.5	0.07
HPC2	2D	*wPt-9997, wPt-4144*	*Xbarc1123*	6.8	0.14
HPC3	3B	*Xbarc0077, Xbarc0229*	*E34M49161L*	6.6	0.16
	3D	*Xcfd70, gpw4125*	*Xbarc1161*	7.2	0.08
VPC1	4D	*Xbarc1118, Xwmc285*	*Xbarc1118*	4.0	0.07
VPC2	-	-	-	-	-
VPC3	1A	*Xbarc0028, E42M49146L*	*CFA2129_RTL*	4.5	0.07
	2B	*Xbarc0328, wPt-9190*	*wPt-2430, Xwmc360*	4.3	0.08
VPC4	6D	*Xbarc0096, Xcfd01*	*Xcfd37*	9.3	0.20
Thousand Kernel Weight	4D	*Xbarc0217, Xbarc1118*	*RHT-DF-MR2*	6.5	0.14
	6D	*Xbarc0328, wPt-9190*	*Xwmc360*	5.2	0.10
	7D	*Xwmc150b, Xbarc172*	*Xwmc150b*	4.6	0.09

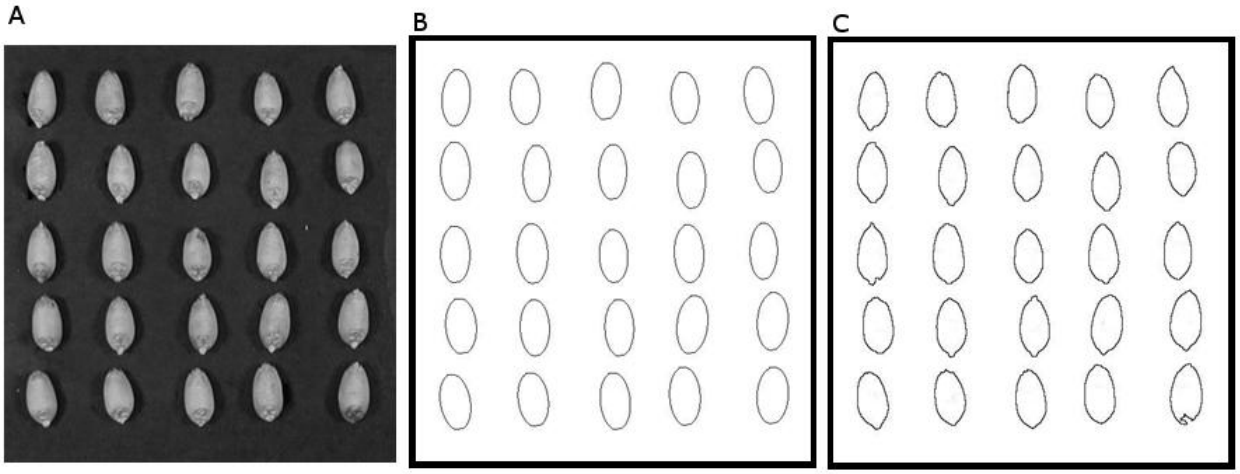

Figure 3.1. Horizontal image (H image) processing via ImageJ with original photograph pictured in A, subsequent transformation to fitted ellipses in B, and outlines used in checking for poor quality measurements shown in C.

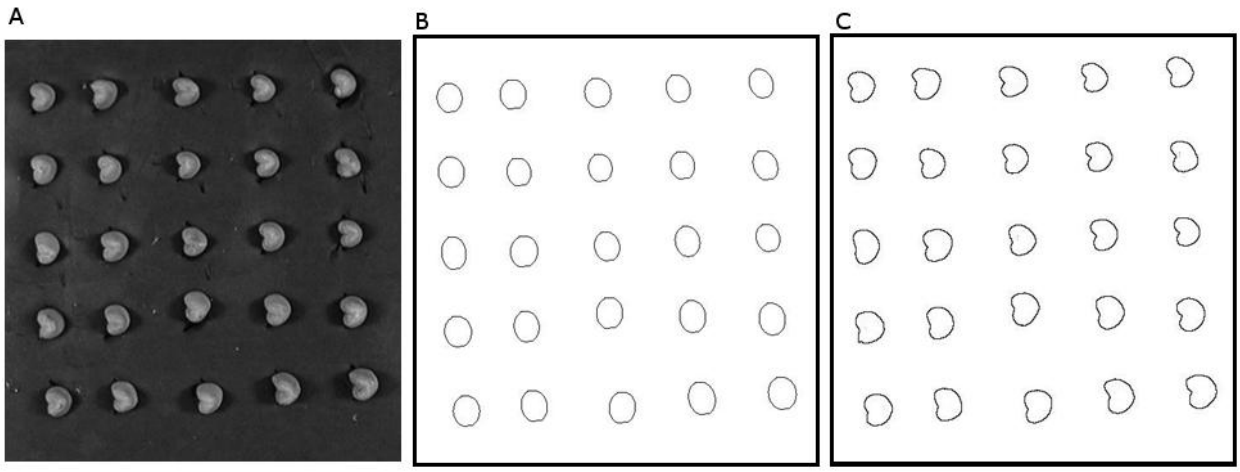

Figure 3.2. Vertical image (V image) processing via ImageJ with original photograph pictured in A, subsequent transformation to fitted ellipses in B, and outlines used in checking for poor quality measurements shown in C.

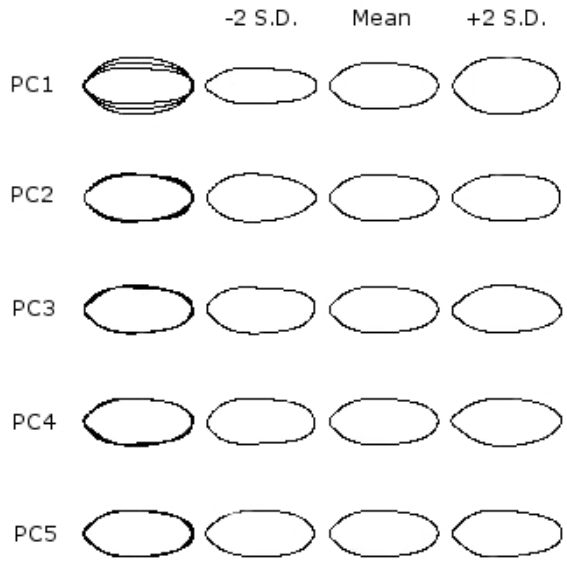

Figure 3.3. Principle components returned from SHAPE for horizontal images of SynOpDH seeds.

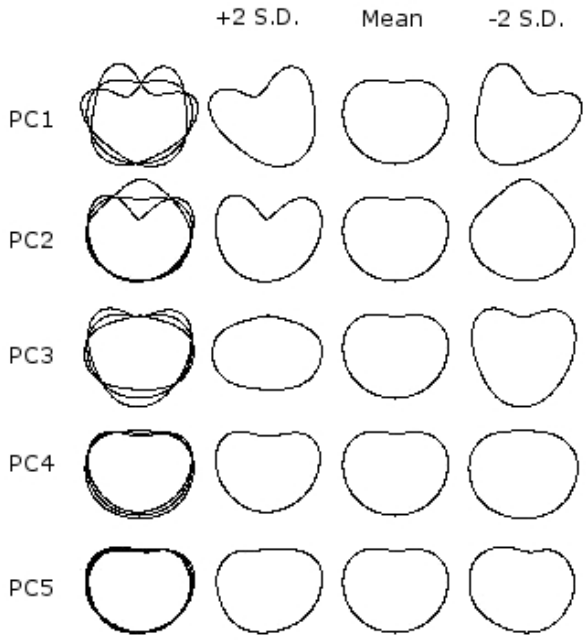

Figure 3.4. Principle components returned from SHAPE for vertical images of SynOpDH seeds.

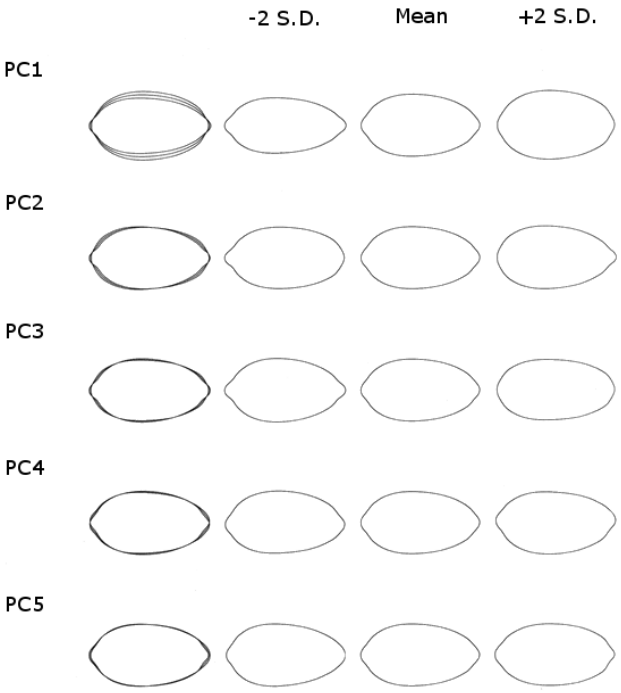

Figure 3.5. Principle components returned from SHAPE for horizontal images of Cayuga X Caledonia seeds.

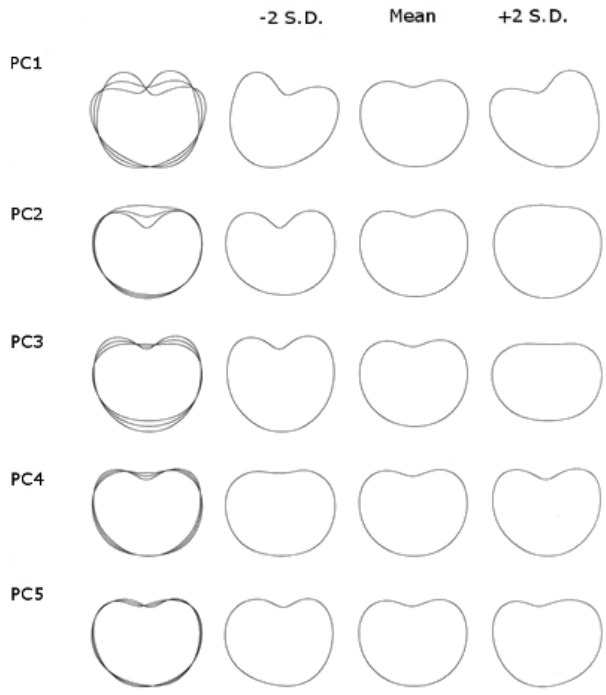

Figure 3.6. Principle components returned from SHAPE for vertical images of Cayuga x Caledonia seeds.

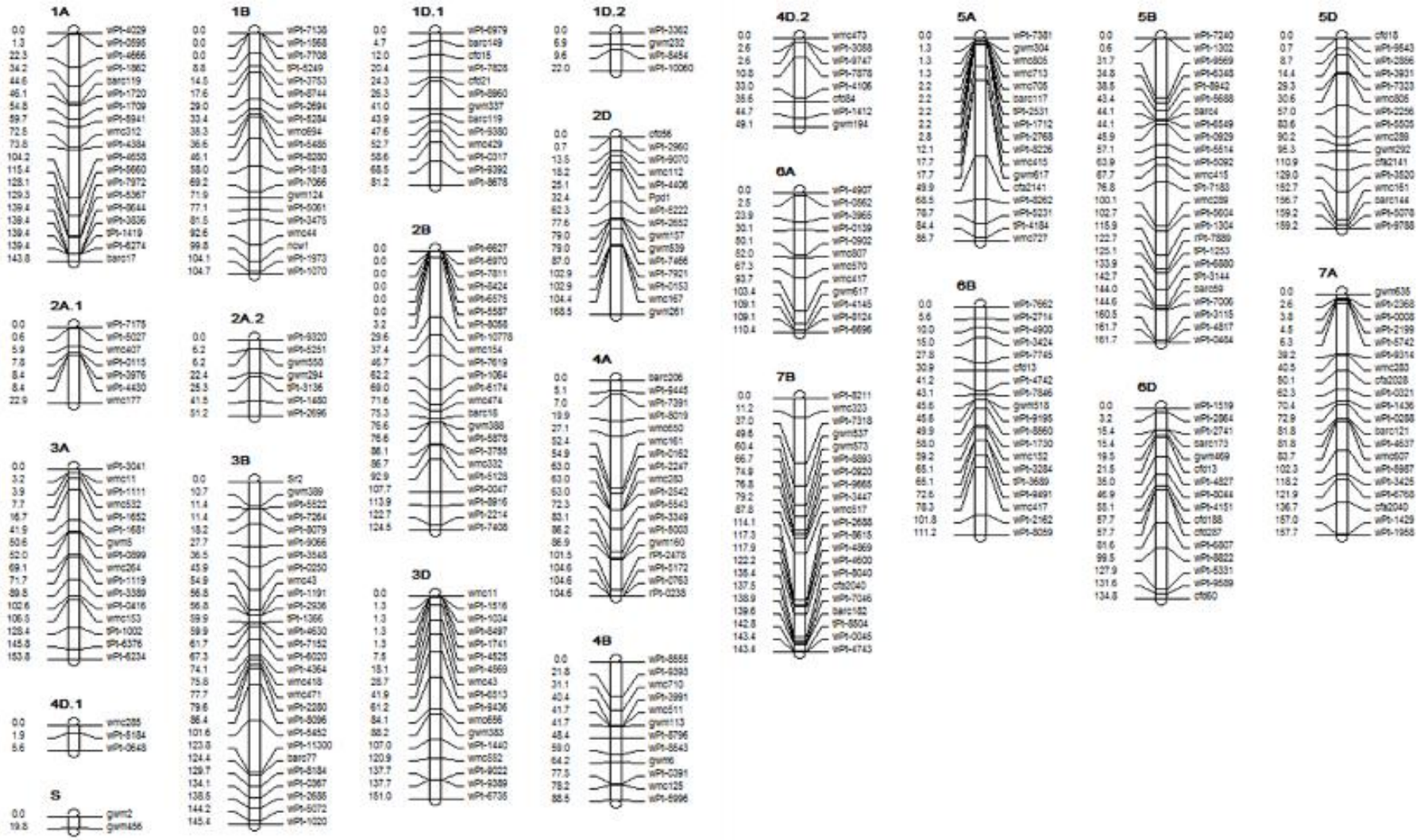

Figure 3.7. Linkage map generated for SynOpDH using reduced marker set

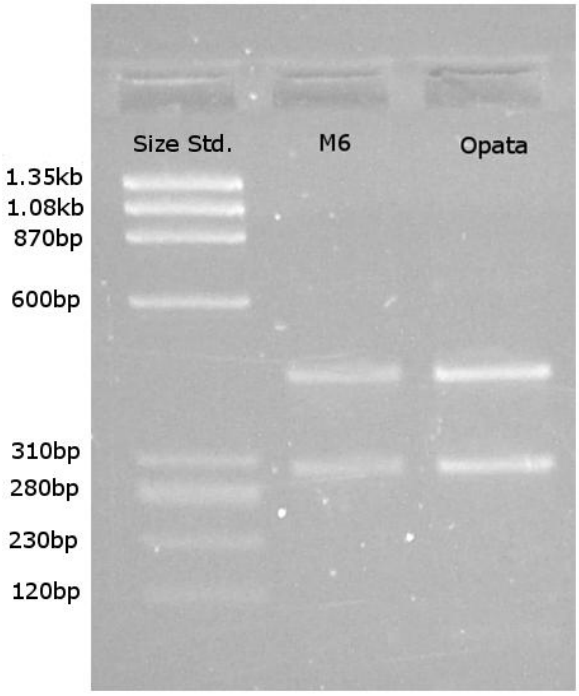

Figure 3.8. Screening of *Ser5B* marker in Synthetic W7984 and Opata M85

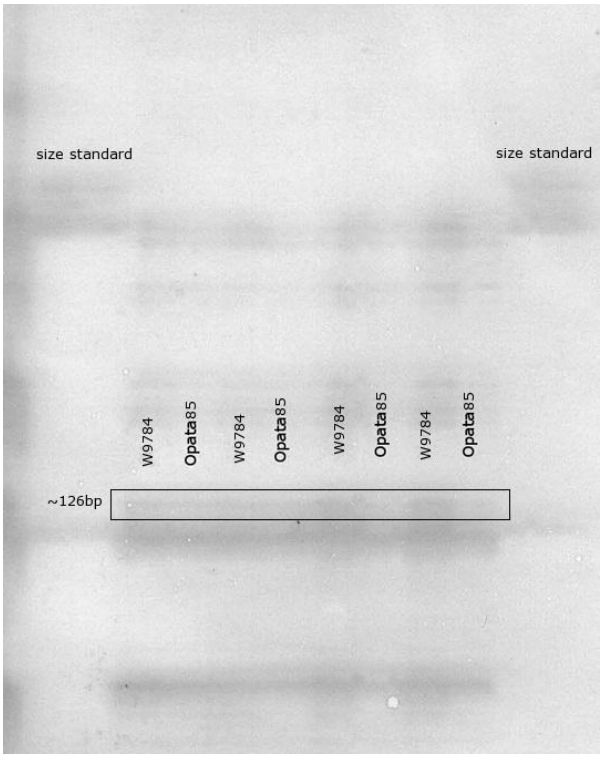

Figure 3.9. Screening of *Gpc-B1* marker in Synthetic W7984 and Opata M85

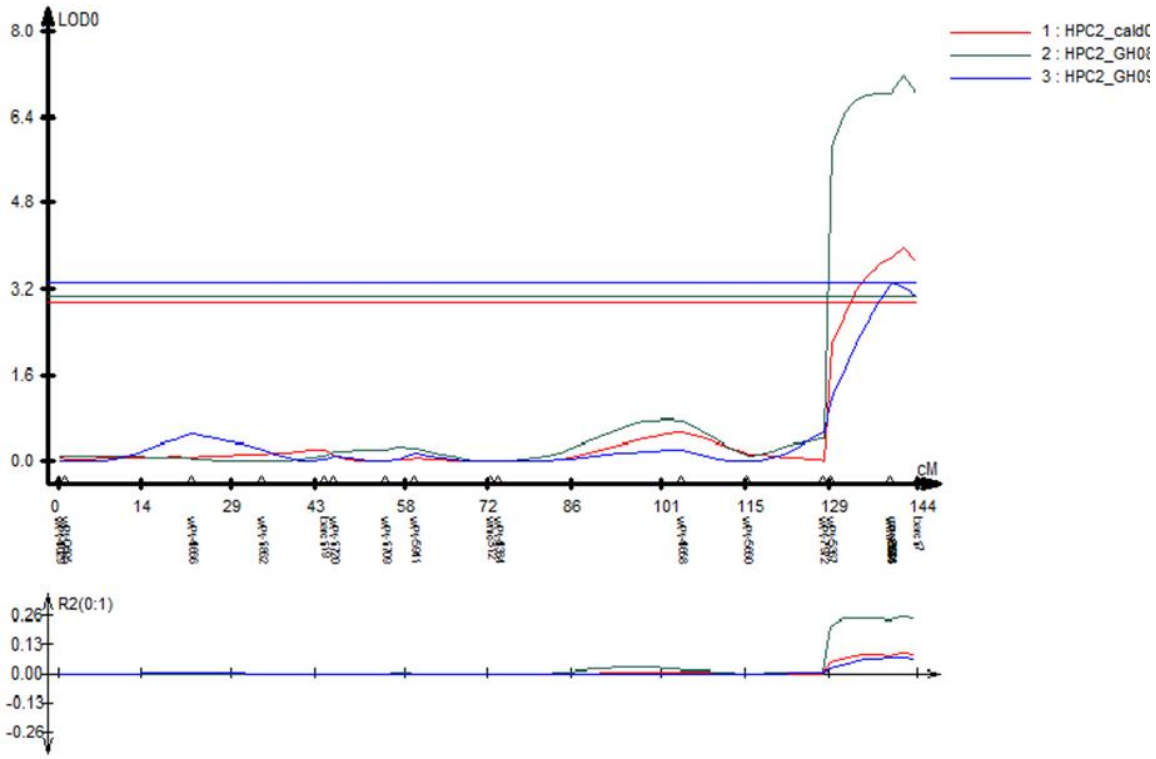

Figure 3.10. QTL detected on chromosome 1A.

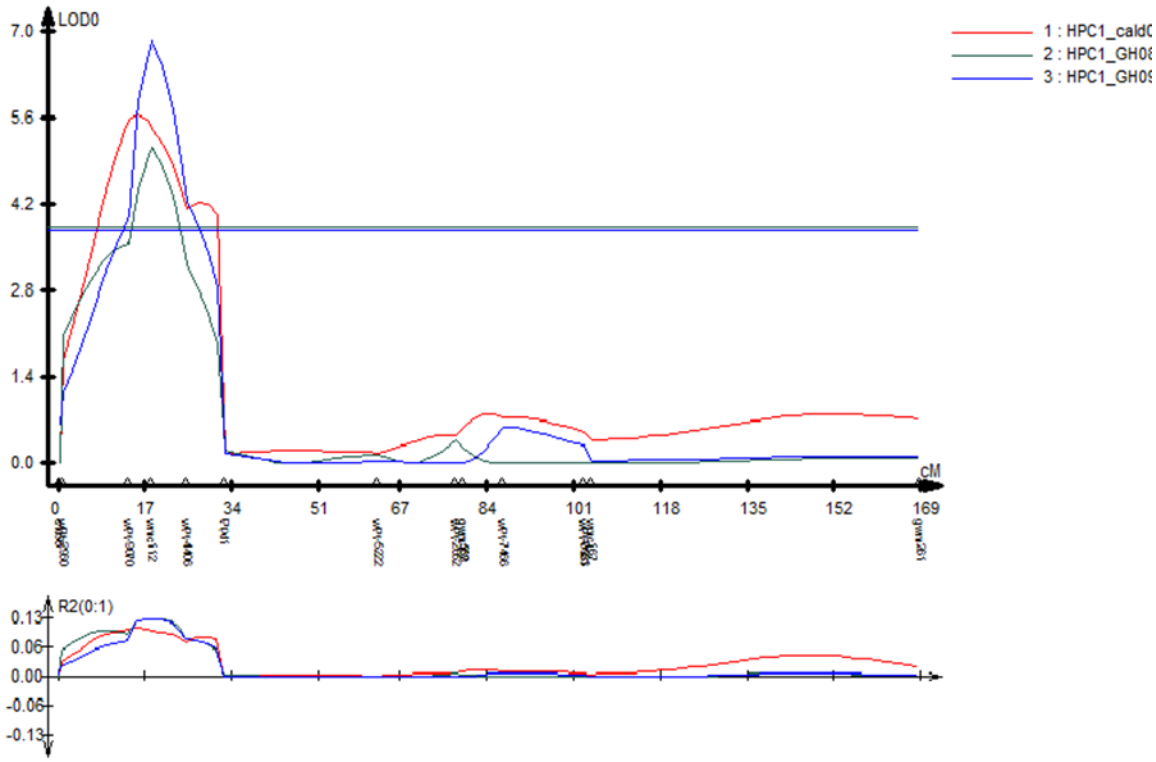

Figure 3.11. QTL detected on chromosome 2D.

REFERENCES

Asakura N, Mori N, Nakamura C, Ohtsuka I (2009) Genotyping of the Q locus in wheat by a simple PCR-RFLP method. Genes Genet Syst 84:233-237

Bergman CJ, Gualberto DG, Campbell KG, Sorrells ME, Finney PL (2000) Kernel morphology variation in a population derived from a soft by hard wheat cross and associations with end-use quality traits. J Food Qual 23:391-407

Breseghello F, Finney PL, Gaines C, Andres L, Tanaka J, Penner G, Sorrells ME (2005) Genetic loci related to kernel quality differences between a soft and a hard wheat cultivar. Crop Sci 45:1685-1695

Breseghello F, Sorrells ME (2006) Association Mapping of Kernel Size and Milling Quality in Wheat (*Triticum aestivum* L.) Cultivars. Genetics 172:1165-1177

Breseghello F, Sorrells ME (2007) QTL analysis of kernel size and shape in two hexaploid wheat mapping populations. Field Crops Research 101:172-179

Campbell KG, Finney PL, Bergman CJ, Gualberto DG, Anderson JA, Giroux MJ, Siritunga D, Zhu J, Gendre F, Roue C, Verel A, Sorrells ME (2001) Quantitative trait loci associated with milling and baking quality in a soft x hard wheat cross. Crop Sci 41:1275-1285

Cane K, Sharp PJ, Eagles HA, Eastwood RF, Hollamby GJ, Kuchel H, Lu M, Martin PJ (2008) The effects on grain quality traits of a grain serpin protein and the VPM1 segment in southern Australian wheat breeding. Australian Journal of Agricultural Research 59:883-890

Cochrane MP (1983) Morphology of the crease region in relation to assimilate uptake and water loss during caryopsis development in barley and wheat. Aust. J. Plant Physiol. 10:473-491

Distelfield A, Uauy C, Fahima T, Dubcovsky J (2006) Physical map of the wheat high-protein content gene *Gpc-B1* and development of a high-throughput molecular marker. New Phytologist 169:753-763

Drea S, Leader DJ, Arnold BC, Shaw P, Dolan L, Doonan JH (2005) Systematic spatial analysis of gene expression during wheat caryopsis development. The Plant Cell 17:2172-2185

Ellis MH, Rebetzke GJ, Chandler P, Bonnett DG, Spielmeyer W, Richards RA (2004) The effect of different height reducing genes on the early growth of wheat. Funct. Plant Biol. 31:583-589

Engledow FL (1920) The inheritance of glume-length and grain-length in a wheat cross. Journal of Genetics 10:109-134

Fischer RA, Stockman YM (1986) Increased kernel number in Norin 10-derived dwarf wheat: evaluation of the cause. Aust. J. Plant Physiol. 13:767-784.

Fischer RA, Quail KJ (1990) The effect of major dwarfing genes on yield potential in spring wheats. Euphytica 46:51-56.

Flintham JE, Borner A, Worland AJ, Gale MD (1997) Optimizing wheat grain yield: effects of *Rht* (gibberellin-insensitive) dwarfing genes. J. Agric. Sci. 128:11-25

Gegas VC, Nazari A, Griffiths S, Simmonds J, Fish L, Orford S, Sayers L, Doonan JH, Snape JW (2010) A Genetic Framework for Grain Size and Shape Variation in Wheat. Plant Cell 22:1046-1056

Guo L, Ma L, Jiang H, Zeng D, Hu J, Wu L, Gao Z, Zhang G, Qian Q (2009) Genetic analysis and fine mapping of two genes for grain shape and weight in rice. Journal of Integrative Plant Biology 51:45-51

Gupta PK, Langridge P, Mir RR (2010) Marker-assisted wheat breeding: present status and future possibilities. Mol Breeding 26:145-161

Hedden P (2003) The genes of the Green Revolution. Trends Genet. 19:5-9

Hook SCW (1984) Specific weight and wheat quality. Journal of the Science of Food and Agriculture 35:1136-1141

Huang XQ, Cloutier S, Lyear L, Radovanovic N, Humphreys DG, Noll JS, Somers DJ, Brown PD (2006) Molecular detection of QTLs for agronomic and quality traits in a doubled haploid population derived from two Canadian wheats (*Triticum aestivum* L.) Theor Appl Genetics 113:753-766

Kang H, Cho Y, Yoon U, Eun M (1998) A rapid DNA extraction method for RFLP and PCR analysis from a single dry seed. Plant Mol Biol Rep 16:90-90

Kato K, Miura H, Sawada S (1999) QTL mapping of genes controlling ear emergence time and plant height on chromosome 5A of wheat. Theoretical and Applied Genetics 98:472-477

Kato K, Miura H, Sawada S (2000) Mapping QTLs controlling grain yield and its components on chromosome 5A of wheat. TAG Theoretical and Applied Genetics 101:1114-1121

Kato K, Sonokawa R, Miura H, Sawada S (2003) Dwarfing effect associated with the threshability gene Q on wheat chromosome 5A. Plant Breeding 122:489-492

Keyes G (1989) *Rht1* and *Rht2* semidwarf genes effect on hybrid vigor and agronomic traits of wheat. Crop Science 29:1442

Kuchel H, Langridge P, Mosionek L, Williams K, Jefferies SP (2006) The genetic control of milling yield, dough rheology and baking quality of wheat. Theor Appl Genet 112:1487-1495

Lamba PS (1949) The relation of glume measurements to kernel shape and size in wheat. Agronomy Journal 41:167

Lehmensiek A, Eckermann PJ, Verbyla AP, Appels R, Sutherland MW, Martin D, Daggard GE (2006) Flour yield QTLs in three Australian doubled haploid wheat populations. Aust. J. of Agric. Research 57:1115-1122

Lingle SE, Chevalier P (1985) Development of the vascular tissue of the wheat and barley caryopsis as related to the rate and duration of grain filling. Crop Sci 25:123-128

Marshall D, Ellison F, Mares D (1984) Effects of grain shape and size on milling yields in wheat. I. Theoretical analysis based on simple geometric models. Australian Journal of Agricultural Research 35:619-630

Marshall D, Mares D, Moss H, Ellison F (1986) Effects of grain shape and size on milling yields in wheat. II. Experimental studies. Australian Journal of Agricultural Research 37:331-342

Maystrenko OI.,Laikova LI, Arbuzova VS, Melnik VM (1998) The chromosomal location of the *S1*, *S2*, and *S3* genes of induced sphaerococcoid mutations in common wheat. EWAC Newsl Proc 10th EWAC meeting, University of Tuscia, Italy:127-130

Millet E, Pinthus MJ (1984) Effects of removing floral organs, light penetration and physical constraint on the development of wheat grains. Annals of Botany 53:261-269

Millet E (1986) Relationships between grain weight and the size of floret cavity in the whet spike. Annals of Botany 58:417-423

Munkvold JD, Tanaka JD, Benscher D, Sorrells ME (2009) Mapping quantitative trait loci for preharvest sprouting resistance in white wheat. Theoretical and Applied Genetics 119:1223-1235

Nalam VJ, Vales MI, Watson CJW, Johnson EB, Riera-Lizarazu O (2007) Map-based analysis of genetic loci on chromosome 2D that affect glume tenacity and threshability, components of the free-threshing habit in common wheat (*Triticum aestivum* L.). Theor Appl Genet 116:135-145

Parker GD, Chalmers KJ, Rathjen AJ, Langridge P (1999) Mapping loci associated with milling yield in wheat (*Triticum aestivum* L.). Molecular Breeding 5:561-568

Pearce S, Saville R, Vaughan SP, Chandler PM, Wilhelm EP, Sparks CA, Al-Kaff N, Korolev A, Boulton MI, Phillips AL, Hedden P, Nicholson P, Thomas SG (2011) Molecular characterization of *Rht-1* dwarfing genes in hexaploid wheat. Plant Physiology 157:1820-1831

Poland JA, Brown PJ, Sorrells ME, Jannink JL (2012) Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS ONE 7(2): e32253

Qiu X, Gong R, Tan Y, Yu S (2012) Mapping and characterization of the major quantitative trait locus *qSS7* associated with increased length and decreased width of rice seeds. Theor Appl Genet 125:1717-1726

Radley M (1981) The effect on wheat grain growth of the removal or ABA treatment of glumes and lemmas. J. Exp. Bot. 32:129-140

Raju GN, Srinivas T (1991) Effect of husk morphology on grain development and topography in rice. Economic Botany 45:429-434

Rasmussen SK, Dahl SW, Norgard A, Hejgaard (1996) A recombinant wheat serpin with inhibitory activity. Plant Molecular Biology 30:673-677

Rebetzke GJ, Richards RA (2000) Gibberellic acid-sensitive dwarfing genes reduce plant height to increase kernel number and grain yield of wheat. Aust. J. Agric. Res. 51:235-245

Rebetzke GJ, Ellis MH, Bonnett D, Mickelson B, Condon AG, Richards, RA (2012) Height reduction and agronomic performance for selected gibberellin-responsive dwarfing genes in bread wheat (*Triticum aestivum* L.). Field Crops Research 126:87-96

Roberts TH, Marttila S, Rasmussen SK, Hejgaard J (2003) Differential gene expression for suicide-substrate serine proteinase inhibitors (serpins) in vegetative and grain tissues of barley. J Exp Bot 54:2251-2263

Rosenkrands I, Hejgaard J, Rasmussen SK, Bjorn SE (1994) Serpins from wheat-grain. FEBS Lett 343:75-80

Salina E, Borner A, Leonova I, Korzun V, Laikova L, Maystrenko O, Roder MS (2000) Microsatellite mapping of the induced sphaerococcoid mutation genes in *Triticum aestivum* . Theoretical and Applied Genetics 100:686-689

Sayre KD, Rajaram S, Fischer RA (1997) Yield potential progress in short bread wheats in northwest Mexico. Crop Sci 37:36-42

Scarth R, Kirby E, Law C (1985) Effects of the photoperiod gene-*Ppd1* and gene-*Ppd2* on growth and development of the shoot apex in wheat. Ann Bot 55:351-359

Schuler SF, Bacon RK, Finney PL, Gbur EE (1995) Relationship of test weight and kernel properties to milling and baking quality in soft red winter-wheat. Crop Sci 35:949-953

Shao G, Wei X, Chen M, Tang S, Luo J, Jiao G, Xie L, Hu P (2012) Allelic variation for a candidate gene for *GS7*, responsible for grain shape in rice. Theor Appl Genet 125:1303-1312

Simons KJ, Fellers JP, Trick HN, Zhang ZC, Tai YS, Gill BS, Faris JD (2006) Molecular characterization of the major wheat domestication gene *Q* . Genetics 172:547-555

Smith N, Guttieri M, Souza E, Shoots J, Sorrels ME, Sneller C (2011) Identification and validation of QTL for grain quality traits in a cross of soft wheat cultivars Pioneer Brand 25R26 and Foster. Crop Sci 51:1424-1436

Somers D, Isaac P, Edwards K (2004) A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.). Theoretical and Applied Genetics 109:1105-1114

Song X, Huang W, Shi M, Zhu M, Lin H (2007) A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. Nat Genet 39:623-630

Sorrells ME, LaRota M, Bermudez-Kandianis CE, Greene RA, Kantety R, Munkvold JD, Miftahudin, Mahmoud A, Ma XF, Gustafson PJ, Qi LLL, Echalier B, Gill BS, Matthews DE, Lazo GR, Chao SM, Anderson OD, Edwards H, Linkiewicz AM, Dubcovsky J, Akhunov ED, Dvorak J, Zhang DS, Nguyen HT, Peng JH, Lapitan NLV, Gonzalez-Hernandez JL, Anderson JA, Hossain K, Kalavacharla V, Kianian SF, Choi DW, Close TJ, Dilbirligi M, Gill KS, Steber C, Walker-Simmons MK, McGuire PE, Qualset CO (2003) Comparative DNA sequence analysis of wheat and rice genomes. Genome Research 13:1818-1827

Sorrells ME, Gustafson PJ, Somers D, Chao S, Benscher D, Guedira-Brown G, Huttner E, Kilian A, McGuire PE, Ross K, Tanaka J, Wenzl P, Williams K, Qualset CO (2011) Reconstruction of the Synthetic W7984 x Opata M85 wheat reference population. Genome 54:875-882

Sourdille P, Tixier MH, Charmet G, Gay G, Cadalen T, Bernard S, Bernard M (2000) Location of genes involved in ear compactness in wheat (*Triticum aestivum*) by means of molecular markers. Mol Breed 6:247-255

Su Z, Hao C, Wang L, Dong Y, Zhang X (2011) Identification and development of a functional marker of TaGW2 associated with grain weight in bread wheat (*Triticum aestivum* L.) Theor Appl Genet 122:211-223

Takano-Kai N, Jiang H, Kubo T, Sweeney M, Matsumoto T, Kanamori H, Padhukasahasram B, Bustamante C, Yoshimura A, Doi K, McCouch S (2009) Evolutionary history of *GS3*, a gene conferring grain length in rice. Genetics 182:1323-1334

Tsilo TJ, Hareland GA, Chao S, Anderson JA (2011) Genetic mapping and QTL analysis of flour color and milling yield related traits using recombinant inbred lines in hard red spring wheat. Crop Sci 51:237-246

Uauy C, Brevis JC, Dubcovsky J (2006) The high grain protein content gene *Gpc-B1* accelerates senescence and has pleiotropic effects on protein content in wheat RID A-4969-2008. J Exp Bot 57:2785-2794

Ugalde TD, Jenner CF (1990a) Route of substrate movement into wheat endosperm. I. Carbohydrates. Aust. J. Plant Physiol. 17:693-704

Ugalde TD, Jenner CF (1990b) Rout of substrate movement into wheat endosperm. II. Amino acids. Aust. J. Plant Phyisol. 17:705-714

Wang G, Leonard JM, Ross AS, Peterson CJ, Zemetra RS, Campbell KG, Lizarazu OR (2012) Identification of genetic factors controlling kernel hardness and related traits in a recombinant inbred population derived from soft x 'extra soft' wheat (*Triticum aestivum* L.) cross. Theor Appl Genet 124:207-221

Wiersma JJ, Busch RH, Fulcher GG, Hareland GA (2001) Recurrent selection for kernel weight in spring wheat. Crop Science 41:999-1005

Williams K, Munkvold JD, Sorrells ME (2012) Comparison of digital image analysis using elliptic Fourier descriptors and major dimensionsto phenotype seed shape in hexaploid wheat. Euphytica DOI: 10.1007/s10681-012-0783-0

Wilson WA, Harrington SE, Woodman WL, Lee M, Sorrels ME, McCouch SR (1999) Inferences on the genome structure of progenitor maize through comparative analysis of rice, maize and the domesticated panicoids. Genetics 153:453-473

Xue S, Zhang Z, Lin F (2008) A high-density intervarietal map of the wheat genome enriched with markers dereived from expressed sequence tags. Theoretical and Applied Genetics 117:181-189.

Yamazaki WT, Briggle LW (1969) Components of test weight in soft wheat. Crop Science 9:457-459

Zhang LY, Liu DC, Guo XL, Yang WL, Sun JZ, Wang DW, Zhang A (2010) Genomic distribution of quantitative trait loci for yield and yield-related traits in common wheat. Journal of Integrative Plant Biology 62:996-1007

CHAPTER FOUR

Effects of selection for seed fill duration on wheat seed shape

Abstract

Wheat seed size and shape potentially impacts quality and yield of wheat cultivars, however their relationship to other traits is complex and poorly understood. The purpose of this study was to characterize seed morphology of the Seed Fill Panel (SFP). The SFP is derived from recurrent selection populations divergently selected for seed fill period length. Image analysis characterized seed morphology in subsets of SFP from either the long or short seed fill progenitor populations. Individual lines had significant variation for all measurements except derived estimates of elliptical shape. Seed characteristics ranked by decreasing phenotypic variation (reported as F-value) were length to width ratio, thickness, length, derived volume, and width. Lines from the long seed fill (LSF) founder population varied more for shape parameters than those from the short seed fill group (SSF). Length of seeds best differentiated subsets, with the SSF material generally having slightly longer seeds. Based on this information and linkage block structure in SFP, it could be used for association mapping of genes underpinning the observed relationship between grain fill and seed morphology.

I. Introduction

Both the rate and the duration of grain filling contribute to the yield of grain crops (Egli 2004). Grain filling and wheat yield relationships were known by the 1970s (Nass and Reiser 1975); and heritable genetic variation exists for both rate and duration of grain filling (Mou and Kronstad 1994a, Wang *et al.* 2009). Grain fill period has been modified as an unintended consequence of selection for yield and earlier flowering by wheat breeding programs (Motzo et al. 2010; Sayre et al. 1997). The rate of grain filling has been implicated in wheat kernel morphology, impacting seed size by regulating

endosperm cell number and resultant sink capacity (Brocklehurst 1977; Mou and Kronstad 1994b). Genetic regions influencing grain filling rate also impact thousand-kernel weight, or seed size (Wang et al. 2009). Furthermore, grain filling characteristics involve source/sink dynamics and yield of wheat is limited by sink capacity rather than assimilate availability (Chojecki *et al.* 1983; Slafer and Savin 1994). Diverse mapping populations and analysis techniques will be needed to fully understand these traits. Quantitative trait loci (QTL) studies have mapped regions of the genome influencing seed morphology using biparental populations (Campbell et al. 1999; Breseghello and Sorrells 2007; Gegas et al. 2009; Sun et al. 2009; Wang et al. 2009; Ramya et al. 2010; Tsilo et al. 2010). However, the QTL and allelic effects are often population specific and difficult to effectively transfer to breeding programs (Stuber et al. 1999; Gupta et al. 2010). Mapping in breeding material,or 'Mapping-As-You-Go' (Podlich *et al.* 2004) can determine effects in a target population, but suffers from limited allelic sampling and reduced genotypic and phenotypic variation (Breseghello and Sorrells 2006). Other genetic mapping approaches such as association mapping (AM), or genome-wide association studies (GWAS), have the potential for increased QTL resolution and allelic sampling although increased resolution depends on patterns of linkage disequilibrium in the mapping panel (Flint-Garcia *et al.* 2003; Breseghello and Sorrells 2006; Rafalski 2010). All of the mapping approaches listed above will help clarify the interrelated genetics of seed filling, morphology, and yield.

Continued development and characterization of populations with variation for seed shape and size is needed to decipher the genetics of these traits. There are many biparental mapping populations and AM panels of existing breeding materials can easily be constructed. What would be most useful is a panel of lines related to breeding material with a high rate of linkage disequilibrium (LD) decay that could allow researchers to fine-map genes of interest. The Seed Fill Panel (SFP) developed at Cornell may be suitable based on a previous report of increased rate of LD decay (Heffner *et al.* 2008). The SFP is derived from recurrent selection populations that were divergently selected for duration of the seed

fill period and lines appear to vary for seed morphology. The objectives of this study were to characterize the phenotypic variation for seed morphology in this panel and report on the heritability of the different seed shape components in the SFP. Divergent selection for seed fill period had the greatest impact on length of seeds, and that the SFP is well suited to future AM studies.

II. Materials and Methods

Germplasm

The seed fill panel (SFP) population consists of 239 recombinant inbred lines derived by single seed descent from two wheat populations segregating for a dominant male sterility gene introduced from the variety Chris hard red spring wheat (Sorrells and Fritz 1982). The founder lines included eight soft white and 8 soft red winter wheat cultivars and 10 soft white and 8 soft red winter wheat experimental lines representative of Eastern United States soft wheat germplasm. They were selected divergently for differences in the length of their grain filling period. The lines derived from the outcrossing population selected for long seed fill (LSF) period are designated with the prefix 'LSF'. Lines derived from the population selected for short seed fill (SSF) period are designated with the prefix 'SSF'. Detailed description of these lines and their potential uses can be found in Heffner *et al.* (2008). A subset of 202 lines with measurements for all seed morphology traits included was used to create a balanced data set for this study and designated 'SSF + LSF'. Two additional subsets were also created, consisting of SSF lines only or LSF lines only.

Growing Environments and Experimental Design

The SFP population was grown in eight environments near Ithaca, NY on the Cornell Research Farm. One location (McGowan) was grown in 2007, three (Snyder, Helfer, and Ketola) were grown in 2008, and four (Caldwell, Helfer, Ketola, and Snyder) were grown in 2009. Environments are

abbreviated using the first three letters of the field name followed by a year number (Mcg07, Sny08, Hel08, Ket08, Cal09, Hel09, Ket09, Sny09). A RCB design was used with two replicates of each entry in Mcg07 but a single replication was used for data analysis because of missing data from the second replicate. In subsequent years (Mcg08, Mcg09) a single entry of each experimental genotype was grown in an augmented design including six check varieties with five replicates at each location. Entries were grown in single one-meter rows, hand harvested, and threshed using a mechanical belt thresher.

Phenotyping Methodology

Phenotypic data was collected using ImageJ from NIH (National Institutes of Health, USA, http://rsbweb.nih.gov/ij/). Seed photography was adapted from the work of Breseghello and Sorrells (2007). The two photographs included a view of the kernel with its crease side down, denoted as the horizontal image or 'H image', and a view of the kernel positioned with the embryo end embedded in the clay, denoted the vertical image or 'V image' (See Figures 4.1 and 4.2). Conversion of images to quantitative measures using ImageJ was adapted from the work of Williams *et al.* (2012). Seed measures were analyzed using analysis of variance (ANOVA) techniques to describe the variety and significance of differences in seed shape. Data was analyzed for the population as a whole as well as for subsets consisting of lines derived either from the LSF or the SSF populations.

Phenotyping shape using axes and pixel counts: The 'Count Object' command of ImageJ returned values for four measures including major axis, minor axis, area, and perimeter of each seed (Figure 4.3). In H images the major axis corresponded to seed length and the minor axis to seed width. In V images the major axis corresponding to seed width and minor axis to seed thickness. Assignment of axes from the separate images to appropriate dimensions was checked by comparison, where width of individual seeds remained consistent between both sets of images (Figure 4.3). Seed images with poor outlines

were removed based on visual inspection and remaining measures were averaged to return phenotypic values for each genotype in each environment.

Phenotyping shape using derived measures: Four derived geometric measures of seed shape were computed (Table 4.1). The volume of seeds was approximated as VOL_{XYZ} using the formula for volume of an ellipsoid (Eric W. Weisstein, MathWorld: http://mathworld.wolfram.com/Ellipsoid.html) based on *x*, *y*, and *z* axes corresponding to seed width, length, and thickness (respectively):

$$\mathrm{VOL_{XYZ}} = \left(\frac{4}{3}\right)\pi xyz$$

The deviation of an individual seed from an optimal ellipse was calculated based on the major and minor axes of either the horizontal image (PDEVH) or the vertical image (PDEVV) (Eric W. Weisstein, Ellipsoid, From MathWorld: http://mathworld.wolfram.com/Ellipsoid.html).

$$p \approx \pi\left[3(a+b) - \sqrt{(3a+b)(a+3b)}\right]$$

where,

p = optimal perimeter value of ellipse

a = major axis of seed

b = minor axis of seed

From *p*, the 'optimal' perimeter value, the actual perimeter measurement from ImageJ was subtracted and divided by the actual perimeter to normalize for differences in length of perimeter (seed size). The

absolute value returned a positive number quantifying how closely the perimeter matched smooth elliptical seed. For example, seeds with rough surfaces returned higher values than smooth seeds.

$$\text{PDEVH} = |(p - \text{HPERIM})/\text{HPERIM}|$$

$$\text{PDEVV} = |(p - \text{VPERIM})/\text{VPERIM}|$$

Analysis of Variance (ANOVA) & Heritability Calculations

Analysis of variance was performed after using several approaches to sub-setting the raw phenotypic data to determine if genotypes possessed variation for the seed characteristics measured (Table 4.2). From ANOVA tables, genotype and environmental mean squares were used to estimate broad-sense heritability as:

$$h^2 = 1\ [\frac{M2}{M1}]$$

Where *M*1 and *M*2 are the mean squares of genotypes and environments, respectively (Huang *et al.* 2006) for each seed characteristic measured (Table 4.3).

Analysis included ANOVA on the 'SSF + LSF' in six environments (Hel09, Ket09, Sny09, Ket08, Sny08, and Mcg07) to provide a completely balanced phenotypic data set. Results from this analysis are reported for only these six environments.

III. Results

Summaries of ANOVA for subsets and all seed characteristics are listed in Table 4.2. Genotypes of the SFP displayed significant variation for all seed characteristics in at least one subset except the

derived measures of PDEVV and PDEVH (Table 4.2). The traits WIDTH and VOL_{xyz} were not significant in the SSF subset.

Genetic Variation in the SSF + LSF subset

When both subsets were included (SSF + LSF) for ANOVA, high levels of variation among lines were observed for seed characteristics (Table 4.2). The most variation was in ASPECT, which had a p-value of 3.27E-75. Seed characteristics ranked by decreasing F-value were ASPECT, THICK, LENGTH, VOL_{xyz}, and WIDTH. The traits PDEVV and PDEVH were not statistically significant among lines.

Seed characteristics in the SSF

The most variation exhibited for any one trait due to genotype was ASPECT, which had a P-value of 6.05E-38. The least variable statistically significant trait was THICK which had a P-value of 0.014692. Individual seed characteristics in the SSF subset ranked by decreasing F-value as ASPECT, LENGTH, and THICK (Table 4.2). The traits WIDTH, VOL_{xyz}, PDEVV, and PDEVH were not statistically significant. Based on the average of all lines in the SSF subset, SSF genotypes generally were slightly larger in individual kernel dimensions than the LSF subset and LENGTH had the largest difference in between subsets (Table 4.4).

Seed characteristics in the LSF

ASPECT had the greatest variation for any one trait due to genotype (P-value = 7.27E-41). The least variable, statistically significant trait was WIDTH (P-value = 2.6E-05), followed by LENGTH (P-value = 2.6E-05). Individual seed characteristics in the LSF subset ranked by decreasing F-value as ASPECT, VOLxyz, THICK, LENGTH, and WIDTH (Table 4.2). The traits PDEVV and PDEVH were not statistically

significant for genotypes. Based on the average of all lines in the LSF subset, LSF genotypes generally were slightly smaller in individual kernel dimensions than the SSF subset (Table 4.4).

Environmental effects

Based on the replicated checks, all environments except Cal09 had coefficients of variation (CV) values < 10% for all direct measures of seed shape (LENGTH, WIDTH, THICK). From the eight initial environments, six were selected for ANOVA based on having a completely balanced data set (Table 4.5). From ANOVA, values for variance of each trait among the 202 lines of 'LSF + SSF' within each environment were calculated. Of direct measures, length displayed the most variability in the different environments, with phenotypic variance among genotypes due to environment ranging from a low value of 0.001367 (Ket08) to a high value of 0.067093 (Sny09). Thickness was least impacted by environment, and displayed the least variability among environments (4.09E-06). Of derived measures, ASPECT displayed the least amount of phenotypic variation due to environment and PDEVH had the greatest differences in phenotypic variation depending on the specific environment. The Ket08 & Ket09 environments had the least impact on phenotypic variance.

IV. Discussion

It is known that seed filling characteristics impact yield and believed that they influence seed morphology as well. Previous selection of seed fill period affected kernel morphology in the SFP. Significant variation for all major seed dimensions in 'SSF + LSF' and the 'LSF' subset was present among all lines when tested by ANOVA. Direct measures of seed shape in the LSF subset had more variation as indicated by smaller p-values for all direct measures of seed shape. The only major dimension that did not display significant variation was WIDTH in the 'SSF' subset. Derived measures VOLxyz and ASPECT were also significant, though PDEVV and PDEVH were not.

Kernel morphology could be used to differentiate the lines derived from either the SSF or LSF subsets. The characteristic that best differentiated lines in either the LSF or SSF subsets was ASPECT, or length -to-width ratio. ASPECT had the smallest p-value in both the SSF and LSF subsets. ASPECT also had the smallest p-value when both subsets were analyzed together (LSF + SSF, p-value of 3.27E-75). The low p-value indicates that ASPECT is the seed characteristic with the largest difference between the two subsets. Given that the trait ASPECT is a derived measure integrating both LENGTH & WIDTH, but WIDTH had much higher p-values, LENGTH is the one dimension most affected by selection based on length of seed fill period.

This shape variation was genetic and could be subject to selection or more detailed mapping studies. Heritability calculations for all seed characteristics in the SFP 'SSF + LSF' subset are higher than those previously reported for biparental mapping populations (Table 4.3). Other studies found heritability of seed characteristics such as thousand-kernel weight ranging from broad-sense heritability of 0.58 to 0.90 and shape parameters (length, width) ranging from 0.55 to 0.95 (Barnard *et al.* 2002; Sun *et al.* 2009; Wang *et al.* 2009; Gegas *et al.* 2010; Tsilo *et al.* 2010). When the SSF and LSF subsets were analyzed individually, heritability values were similar to the 'SSF + LSF' subset for all traits.

The phenotypic variation and unique recombination history of the SFP make it useful for association mapping. Reduced linkage blocks (increased LD decay) exist in the SFP due to the dominant male-sterile progenitors of the SFP undergoing annual cycles of intermating (Heffner et al. 2008). Increased LD decay is atypical of autogamous crops like wheat. Increased LD decay means that positive marker trait associations detected using SFP are more likely from the marker being in tight linkage with a functional polymorphism rather than being identified as artifacts of population structure. Presumably there are multiple alleles for seed shape loci in SFP (versus two in a biparental population) that were originally present in adapted cultivars. These alleles may be more useful for breeding programs since they are functional in a panel derived from cultivars (Breseghello and Sorrells 2006). Given increased rate of LD

decay and phenotypic variation described here, with sufficient marker density it is likely that markers linked to the functional polymorphisms impacting seed shape and grain fill characteristics can be identified in the SFP. The SFP could then be used for fine-mapping, cloning, or verification of these alleles.

The SFP represents a unique source of phenotypic variation for seed morphology studies in hexaploid wheat. The variation is derived from alleles affecting seed fill and kernel morphology in adapted wheat cultivars and represents a wider allelic sampling than occurs in biparental mapping or breeding populations. The SFP can elucidate how selection for grain fill period impacts individual seed dimensions, as suggested by the difference in seed length between the SSF and LSF subsets. In addition to the SFP itself, the image files represent a resource for exploring different phenotyping approaches using photometrics. Previous empirical studies describing LD decay in SFP (Heffner et al. 2008) suggest that it is a powerful tool for association mapping of these traits. It is hoped that future genotyping of the SFP provides the data needed to extend this phenotypic characterization to a formal mapping study.

Table 4.1. Descriptions of seed morphology phenotypes.

Direct Photometric Measures, ImageJ	**Abbreviation**	**Measured as:**
Seed Length	LENGTH	Major axis of horizontal image
Seed Width	WIDTH	Minor axis of horizontal image (equivalent to major axis of vertical image)
Seed Thickness	THICK	Minor axis of vertical image
Derived Photometric Measures	**Abbreviation**	**Derivation:**
Aspect Ratio	ASPECT	(LENGTH) / (WIDTH)
Volume	VOL_{XYZ}	(4/3)π(x)(y)(z), where x, y, & z represent LENGTH, WIDTH, & THICK
Deviation from an optimal ellipse, horizontal	PDEVH	PDEVH = |(p – HPERIM) / HPERIM | where p = π [3(LENGTH + WIDTH) – √ (3*LENGTH + WIDTH)*(LENGTH + 3*WIDTH)
Deviation from an optimal ellipse, vertical	PDEVV	PDEVV = |(p – VPERIM) / VPERIM | where p = π [3(THICK + WIDTH) – √ (3*THICK + WIDTH)*(THICK + 3*WIDTH)

Table 4.2. ANOVA tests for variation in seed morphology traits among the SFP panel including the balanced complete panel (SSF + LSF) and balanced subsets of short seed fill (SSF) and long seed fill (LSF) derived lines; traits which do not reject the null hypothesis of no variation (based on a significance of <0.05) are highlighted.

	SSF + LSF			**SSF**			**LSF**		
	F	*F crit*	*P-value*	*F*	*F crit*	*P-value*	*F*	*F crit*	*P-value*
LENGTH	1.495974	1.189846	5.38E-05	1.422534	1.261572	0.006044	2.052489	1.291419	8.76E-07
WIDTH	1.26716	1.189846	0.012327	1.201948	1.261572	0.09648	1.848195	1.291419	2.6E-05
THICK	1.502313	1.189846	4.35E-05	1.359052	1.261572	0.014692	2.30712	1.291419	9.89E-09
VOL_{xyz}	1.335632	1.189846	0.002923	1.232211	1.261572	0.069731	2.864481	1.291419	3.08E-13
ASPECT	5.533329	1.189846	3.27E-75	5.08735	1.261572	6.05E-38	6.332786	1.291419	7.27E-41
PDEVV	1.003753	1.189846	0.476648	1.086228	1.261572	0.27513	0.870916	1.291419	0.786991
PDEVH	0.890719	1.189846	0.846276	0.880526	1.261572	0.792612	1.000108	1.291419	0.485251

Table 4.3. Heritability for seed morphology traits among the SFP panel including the balanced complete panel (SSF + LSF) and subsets of short seed fill (SSF) and long seed fill (LSF) derived lines

	SSF + LSF	SSF	LSF
LENGTH	0.99	0.99	0.99
WIDTH	0.99	0.99	0.99
THICK	0.99	0.99	0.99
VOL_{xyz}	0.99	0.99	0.99
ASPECT	0.99	0.99	0.99
PDEVV	0.99	0.99	0.99
PDEVH	0.99	0.99	0.99

Table 4.4. Average phenotypic LSF and SSF subset values of each primary dimension of the seed in each environment sampled

		LENGTH (cm)	WIDTH (cm)	THICK (cm)
Hel09	**LSF**	0.71847	0.26445	0.279356
	SSF	0.723579	0.267158	0.284884
Ket09	**LSF**	0.723008	0.353306	0.25416
	SSF	0.73272	0.35854	0.261048
Sny09	**LSF**	0.723848	0.25991	0.253315
	SSF	0.733304	0.262713	0.256122
Ket08	**LSF**	0.71303	0.369144	0.293662
	SSF	0.71814	0.36888	0.294183
Sny08	**LSF**	0.684543	0.332715	0.255081
	SSF	0.725671	0.362244	0.269018
Mcg07	**LSF**	0.721177	0.529142	0.215586
	SSF	0.709996	0.532908	0.212852

Table 4.5. Phenotypic variance within environment for each trait in 'LSF + SSF' subset

Environment	Phenotypic Variances						
	LENGTH	WIDTH	THICK	VOL_{xyz}	ASPECT	PDEVV	PDEVH
Hel09	.005269	.002012	.001208	.002514	.038665	.001144	2.95e-7
Ket09	.001553	.000344	.000341	2.79e-5	.019129	.001265	1.77e-7
Sny09	.006132	.000807	.000309	.000105	.04348	.001127	3.87e-7
Ket08	.001367	.00049	.000327	3.46e-5	.019843	.000354	4.42e-7
Sny08	.067093	.038411	.005414	.000339	.021598	.006189	3.645808
Mcg07	.01942	.008963	.002554	.000417	.002991	.014872	2.317529

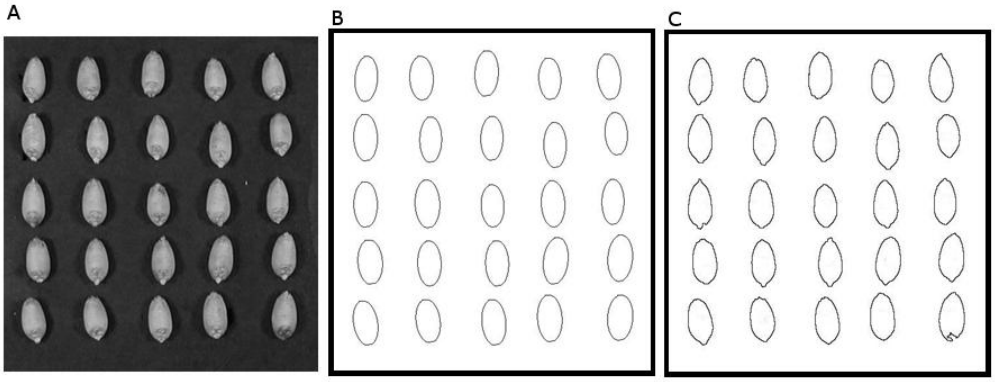

Fig 4.1. Horizontal image (H image) processing via ImageJ with original photograph pictured in A, subsequent transformation to fitted ellipses in B, and outlines used in checking for poor quality measurements shown in C.

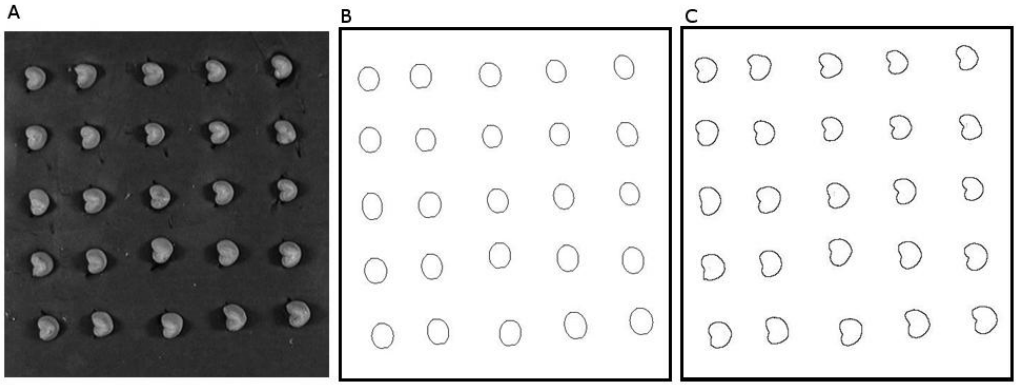

Fig 4.2. Vertical image (V image) processing via ImageJ with original photograph pictured in A, subsequent transformation to fitted ellipses in B, and outlines used in checking for poor quality measurements shown in C.

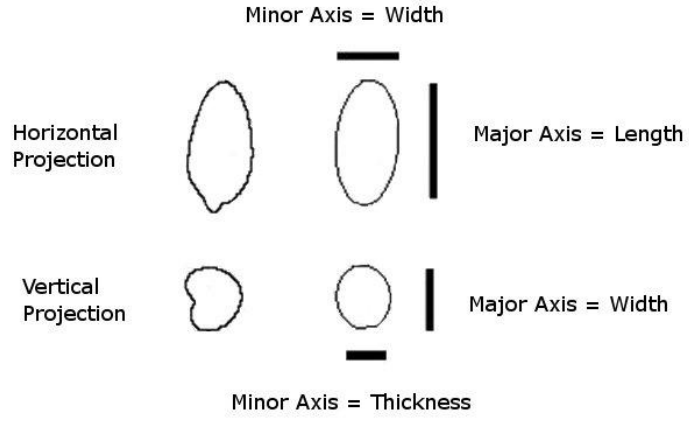

Fig 4.3. Conversion of seed images into axes measurements via ImageJ.

REFERENCES

Barnard AD, Labuschagne MT, van Nierkerk HA (2002) Heritability estimates of bread wheat quality traits in the Western Cape province of South Africa. Euphytica 127:115-122

Breseghello F, and Sorrells ME (2006) Association analysis as a strategy for improvement of quantitative traits in plants. Crop Sci 46:1323-1330

Breseghello F, Sorrells ME (2007) QTL analysis of kernel size and shape in two hexaploid wheat mapping populations. Field Crops Research 101:172-179

Brocklehurst PA (1977) Factors controlling grain weight in wheat. Nature 266:348-349

Campbell KG, Bergman CJ, Gualberto DG, Anderson JA, Giroux MJ, Hareland G, Fulcher RG, Sorrells ME, Finney PL (1999) Quantitative trait loci associated with kernel traits in a soft x hard wheat cross. Crop Sci 39:1184-1195

Chojecki AJS, Gale MD, Bayliss MW (1983) Reciprocal monosomic analysis of grain size in wheat. In: Sakamoto S (Ed) Proc 6th Int Wheat Genet Symp. Maruzen, Kyoto, Japan. 1061

Flint-Garcia SA, Thornsberry JM, Buckler ES (2003) Structure of linkage disequilibrium in plants. Annu. Rev. Plant Biol. 54:357-374

Gegas V, Nazari A, Griffiths S, Simmonds J, Fish L, Orford S, Sayers L, Doonan JH, Snape JW (2010) A genetic framework for grain size and shape and shape variation in wheat. The Plant Cell 22:1046-1056

Gupta PK, Langridge P, Mir RR (2010) Marker-assisted wheat breeding: present status and future possibilities. Mol Breeding 26:145-161

Heffner EL, Chomdej O, Williams KR, Sorrells ME (2008) Dominant male-sterile populations for association mapping and introgression of exotic wheat germplasm. Australian Journal of Agricultural Research 59:470-474

Huang XQ, Cloutier S, Lycar L, Radovanovic N, Humphreys DG, Noll JS, Somers DJ, Brown PD (2006) Molecular detection of QTLs for agronomic and quality traits in a doubled haploid population derived from two Canadian wheats (*Triticum aestivum* L.) Theor Appl Genet 113: 753-766

Motzo R, Francesco G, Giovanni P (2010) The response of rate and duration of grain filling to long-term selection for yield in Italian durum wheats. Crop & Pasture Science 61:162-169

Mou BQ, Kronstad WE (1994a) Duration and rate of grain filling in selected winter-wheat populations. 1. Inheritance. Crop Science 34:833-837

Mou BQ, Kronstad WE (1994b) Grain filling parameters and protein-content in selected winter-wheat populations. 2. Associations. Crop Science 34:838-841

Nass HG, Reiser B (1975) Grain filling period and grain yield relationships in spring wheat. Canadian Journal of Plant Science 3:673-678

Podlich DW, Winkler CR, Cooper M (2004) Mapping as you go: an effective approach for marker-assisted selection of complex traits. Crop Sci 44:1560-1571

Ramya P, Chaubal A, Kulkarni K, Gupta L, Kadbo N, Dhaliwal HS, Chhuneja P, Lagu M, Gupta V (2010) QTL mapping of 1000-kernel weight, kernel length, and kernel width in bread wheat (*Triticum aestivum* L.) J Appl Genet 51:421-429

Rafalski JA (2010) Association genetics in crop improvement. Curr. Opin. Plant Biol. 13:174-180

Sayer KD, Rajaram S, Fischer RA (1997) Yield potential progress in short bread wheats in northwest Mexico. Crop Sci 37:36-42

Sharma RC (1994) Early generation selection for grain filling period in wheat. Crop Science 4:945-948

Slafer GA, Savin R (1994) Source-sink relationships and grain mass at different positions within the spike in wheat. Field Crops Research 37:39-49

Sorrells ME, Fritz SE (1982) Application of a dominant male-sterile allele to the improvement of self-pollinated crops. Crop Science 22:1033-1035

Stuber CW, Polacco M, Lynn SM (1999) Synergy of empirical breeding, marker-assisted selection, and genomics to increase crop yield potential. Crop Sci 40:778-782

Sun XY, Wu K, Zhao Y, Kong FM, Han GZ, Jiang HM, Huang XJ, Li RJ, Wang HG, Li SS (2009) QTL analysis of kernel shape and weight using recombinant inbred lines in wheat. Euphytica 165:615-624

Tsilo TJ, Hareland GA, Simsek S, Chao S, Anderson JA (2010) Genome mapping of kernel characteristics in hard red spring wheat breeding lines. Theor Appl Genet 121:717-730

Wang R, Hai L, Zhang X, You G, Yan C, Xiao S (2009) QTL mapping for grain filling rate and yield-related traits in RILs of the Chinese winter wheat population Heshangmai x Yu8679. Theor Appl Genet 118:313-325

Williams K, Munkvold J, Sorrells M (2012) Comparison of digital image analysis using elliptic Fourier descriptors and major dimensions to phenotype seed shape in hexaploid wheat (*Triticum aestivum* L.) Euphytica DOI 10.10007/s10681-012-0783-0